Cell-Specific Role of Protease 2A in Poliovirus Translation in Relation to 5' Noncoding Region Secondary Structures

Thesis

UNRESTRICTED

Cell-specific role of protease 2A in poliovirus translation in relation to 5' noncoding region secondary structures

by

Alison Rowe ARCS BSc

National Institute for Biological
Standards and Control
South Mimms

A thesis submitted in partial fulfilment
of the requirements of the Open University
for the degree of Doctor of Philosophy
under the discipline of
Life Sciences

August 1997

DATE OF AWARD: 12 JANUARY 1998

CONTENTS PAGE

DECLARATION

I declare that this thesis is my own composition and that the research described herein was performed entirely by myself except where expressly stated.

Alison Rowe
June 1997

ABSTRACT

Polioviruses are single stranded RNA viruses and members of the picornavirus family. They have an unusually long non coding region at the 5' end which has elaborate secondary structure predicted to be made up of seven domains. Part of the secondary structure constitutes an internal ribosome entry site (IRES) which allows the virus to utilise cap-independent initiation of translation whilst disabling the host cell's cap-dependent translation. Translation inefficiencies originally caused by disruptions of the secondary structure of domain V could be compensated for in monkey kidney cells *in vitro* by coding changes in the protease 2A. These coding changes were found throughout the protease which is otherwise quite highly conserved in polioviruses. They appear to have no effect on monkey neurovirulence and their activity was found to be cell specific, having little or no compensatory effect in a mouse cell line. Using this cell line to investigate the effects of mutations in functionally significant secondary structure of domain V on virus growth highlighted unpaired loops as being important. Shortening one loop was detrimental to the virus whereas the sequence of the loop was much less important. In addition, flexibility at another position between two stems appeared to be critical and mutations that could potentially alter the folding in this area rendered the virus temperature sensitive.

This Egyptian funerary stele, dated around 1300BC, probably depicts the earliest record of poliomyelitis. The priest, Rom, shows a withered leg and dropped foot, typical signs of paralytic poliomyelitis. The stele currently resides in the Ny Carlsberg Glyptothek museum in Copenhagen.

LIST OF FIGURES AND TABLES

CHAPTER 4

CHAPTER 5

CHAPTER 6

CHAPTER 7

APPENDIX

ACKNOWLEDGEMENTS

Enormous appreciation goes to Glynis for all her support during my stay in her laboratory. She is an absolute treasure and it was an honour to be able to work with somebody of her experience. I will miss the singing. Appreciation also goes to Ges for all her help over the years, and for her cheeriness. Both Glynis and Ges have been indispensable throughout my time at NIBSC.

Thanks go to Lesley for her help with cells, to Miggy for making up all the medium and various solutions and to Glynis and Carolyn for making the oligos. Thanks also go to Mac for his help with all my computing problems (where was the pentium when I needed it?) and to Ian Feavers for assisting me in the Staden package. All my photographs and some illustrations were by Robert Longuehaye and the basic domain V structures and scanned images were by Andrew Davies. I must also thank Denise Lawrence for her help with references and Phil Yates for lending me his computer.

Thanks go to my three supervisors. Phil Minor was a great help in reading and discussing my thesis. I will always appreciate the fact that he originally let me work in his division all those years ago, when my interest in polioviruses was initiated. I also thank Andrew Macadam for his support during my project and Jim Robertson for his enthusiasm with reading my thesis.

Last but by no means the least I have to thank Robert for his support and nagging. Always there when I needed help, without him I don't know how I'd have ever got through all this.

This thesis is dedicated to Robert.

A ward of 'iron lungs'.
The paralytic effects of poliomyelitis are devastating and permanent. This is the reason why the WHO are so insistent on the eradication of the disease and the causative agent. This is what initiated the many years of research on polioviruses of which this thesis is a part.

ABBREVIATIONS

°C	degrees Celsius
µ	micro, 10^{-6}
cal	calories
CELICS	cloning by enzyme-linked immuno-colour screening
Ci	Curies
CNS	central nervous system
CPE	cytopathic effect
CVI	Children's Vaccine Initiative
Da	Dalton
dATP	deoxy-adenosine-triphosphate
dCTP	deoxy-cytidine-triphosphate
ddATP	dideoxy-adenosine-triphosphate
ddCTP	dideoxy-cytidine-triphosphate
ddGTP	dideoxy-guanosine-triphosphate
ddITP	dideoxy-inosine-triphosphate
ddTTP	dideoxy-thymidine-triphosphate
dGTP	deoxy-guanosine-triphosphate
dITP	deoxy-inosine-triphosphate
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic acid
dNTP	dATP, dGTP, dCTP and/or dTTP
ds	double stranded
DTT	1,4-dithio-threitol
dTTP	deoxy-thymidine-triphosphate
EDTA	ethylenediaminetetraacetic acid
E. Coli	*Escherichia coli*
FAB	fragment antigen binding
FCS	foetal calf serum
Fu	fungizone
G	acceleration in the Earth's magnetic field
g	gramme
GCG	Genetics Computer Group
GDW	glass distilled water
HBSS	Hank's balanced salt solution
HEPES	N-2-hydroxymethylpiperazine-N'-2-ethane sulphonic acid
HIV	Human immunodeficiency virus
IPTG	isopropyl-β-D-thiogalactoside
IRES	internal ribosome entry site
k	kilo, 10^{3}
l	litre
m	milli, 10^{-3}
M	molar
min	minute
mol	mole
mRNA	messenger RNA
NCR	non coding region
NIBSC	National Institute for Biological Standards and Control
p	pico, 10^{-9}

PBS	phosphate buffered saline
PCR	polymerase chain reaction
P/S	penicillin and streptomycin
RNA	ribonucleic acid
s	second
S	Svedberg units
SDS	sodium dodecyl sulphate
ss	single stranded
TEMED	tetramethyl-ethylenediamine
tris	tris(hydroxymethyl)methylamine
tris-HCl	2-amino-2-(hydroxy-methy)-1.3-propanediol-hydrochloride
ts	temperature sensitive
WHO	World Health Organisation
X-Gal	5'-bromo-4-chloro-3-indolyl-β-D-galacto-pyranoside

AMINO ACID ABBREVIATIONS AND UNIVERSAL CODONS

Alanine	Ala	A	GCU, GCC, GCA, GCG
Arginine	Arg	R	AGA, AGG, CGU, CGC, CGA, CGG
Asparagine	Asn	N	AAU, AAC
Aspartic acid	Asp	D	GAU, GAC
Cysteine	Cys	C	UGU, UGC
Glutamic acid	Glu	E	GAA, GAG
Glutamine	Gln	Q	CAA, CAG
Glycine	Gly	G	GGU, GGC, GGA, GGG
Histidine	His	H	CAU, CAC
Isoleucine	Ile	I	AUU, AUC, AUA
Leucine	Leu	L	CUU, CUC, CUA, CUG, UUA, UUG
Lysine	Lys	K	AAA, AAG
Methionine	Met	M	AUG
Phenylalanine	Phe	F	UUU, UUC
Proline	Pro	P	CCU, CCC, CCA, CCG
Serine	Ser	S	UCU, UCC, UCA, UCG, AGU, AGC
Threonine	Thr	T	ACU, ACC, ACA, ACG
Tryptophan	Trp	W	UGG
Tyrosine	Tyr	Y	UAU, UAC
Valine	Val	V	GUU, GUC, GUA, GUG

CHAPTER ONE

INTRODUCTION

1.1 GENERAL INTRODUCTION

Poliomyelitis derives its name from the Greek words 'polios' meaning grey and 'muelos' meaning spinal cord and polioviruses are the causative agents of the disease. It refers to the destruction of the neurones in the grey matter of the spinal cord, a characteristic of the less common but more visible major illness responsible for paralysis. An Egyptian funerary stele dated around 1300BC depicting the priest Rom with a withered limb is probably the earliest record of the disease but very few descriptions were documented until around the turn of this century. In 1988 the WHO announced its intention of totally eradicating poliomyelitis by the year 2000. Although here in Great Britain and developed countries such as the USA the occurrence of poliomyelitis is now extremely rare there still remain a number of places in the world where the disease is common (see fig 1.1) and the WHO's target may need to be extended.

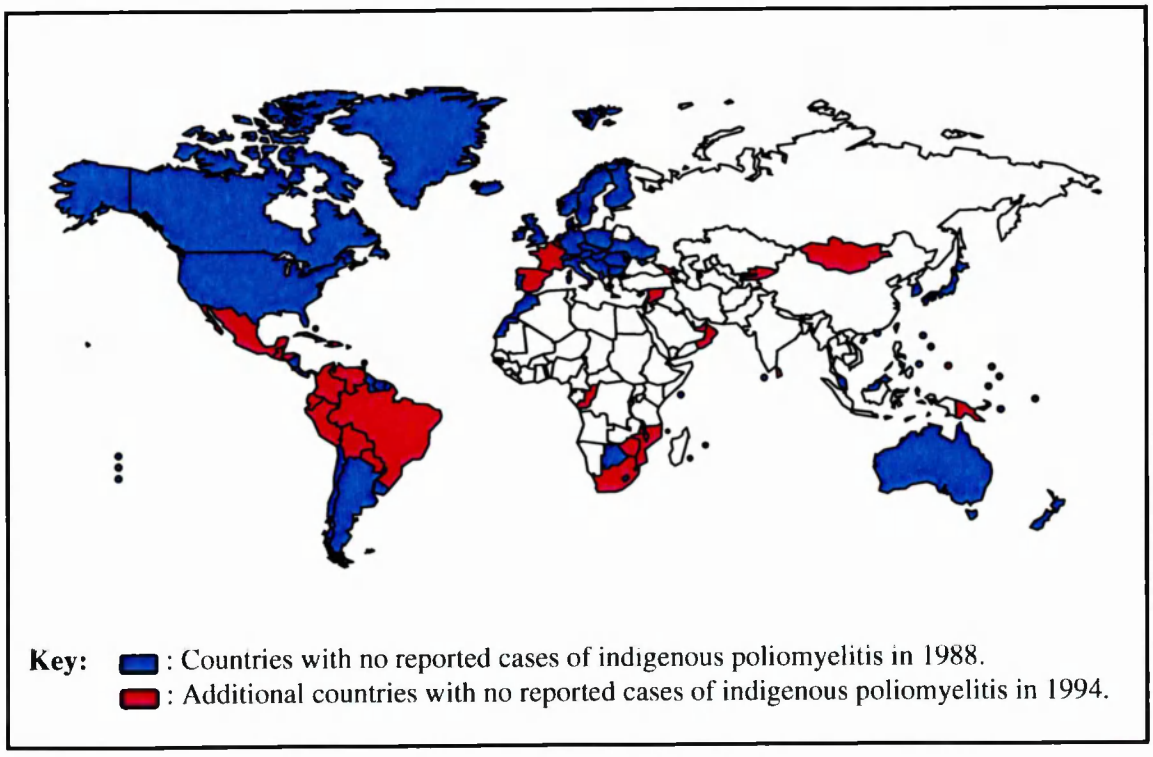


FIGURE 1.1 Countries with no indigenous poliovirus. The countries that had no reports of wild type virus infections by the year 1988 are coloured blue and the additional countries that have no reports of wild type virus infections between 1988 and 1994 are coloured red. Adapted from WHO CVI Forum, No 9, 1995.

Before the introduction of vaccines poliomyelitis was understandably feared, much as AIDS is today. The non reversible paralysis and life in an 'iron lung', although statistically rare complications of poliomyelitis, were very apparent to the general public. Numerous outbreaks in the first half of this century were instrumental in initiating the search for an effective vaccine but funding for this was intensified only in 1933 when F D Roosevelt, severely paralysed by poliomyelitis, became president of the United States. Passionate appeals from the president for sponsorship were successful and enabled first Jonas Salk, then Albert Sabin to formulate the vaccines that are still in use today.

1.2 THE VIRUS

1.2.1 GENERAL

Polioviruses are classified as members of the enterovirus genus of the Picornaviridae family. Currently there are five genera with the suggestion that some echoviruses now form a genus on their own called orphanovirus (fig 1.2). Picornaviruses derive their name from being small (pico) RNA viruses. The family of viruses is implicated in a wide range of diseases but all members share similar genomic and structural features.

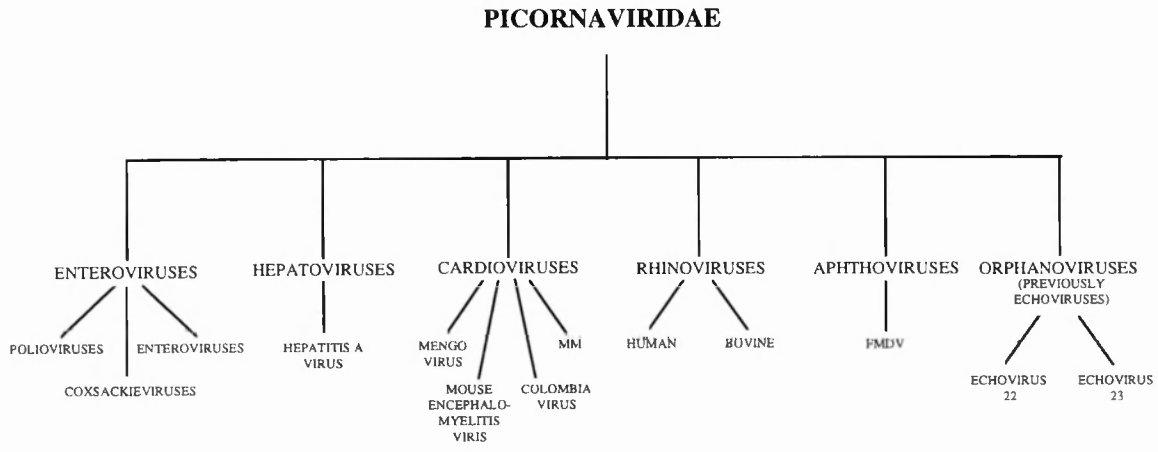


FIGURE 1.2 Diagram of the Picornaviridae family illustrating the five genera: enteroviruses; hepatoviruses; cardioviruses; rhinoviruses; aphthoviruses and the suggested sixth genus, orphanoviruses. Representative viruses of each genus are shown.

Polioviruses are approximately 27nm in diameter, non enveloped and have a single copy of positive strand RNA approximately 7500 nucleotides long. They occur in three serotypes: type 1; type 2 and type 3 with type 1 responsible for most wild type paralytic infections, type 3 being found in a small proportion and type 2 in very few. A number of different strains of virus exist including the 3 attenuated Sabin strains that are used in the oral vaccine and many wild type viruses recovered from human cases of the disease *e.g.* Mahoney (type 1) and Leon (type 3). The first complete viral sequence to be published was of a type 1 Mahoney virus (Kitamura *et al.*, 1981) and many others have followed (Racaniello & Baltimore 1981a; Nomoto *et al.*, 1982; Toyoda *et al.*, 1984; Stanway *et al.*, 1984a; Cann *et al.*, 1984; Hughes *et al.*, 1986; Pollard *et al.*, 1989; reviewed in Stanway 1990). Humans are the only natural host for the virus although some primates can be infected. Polioviruses can be adapted to infect mice (Armstrong 1939a; Armstrong 1939b; Racaniello 1984; La Monica *et al.*, 1986) but once the human poliovirus receptor gene was cloned (Mendelsohn *et al.*, 1989) various transgenic murine lines were made (Ren *et al.*, 1990; Jubelt *et al.*, 1991) and research into their use as animal models of infection is on-going.

1.2.2. STRUCTURE OF THE VIRUS

The viral capsid is made up of 60 copies of each of the virion proteins, VP1, VP2, VP3 and VP4. They form an icosahedral particle (see figs 1.3 & 1.4), whose structure has been determined by x-ray crystallographic methods (Hogle *et al.*, 1985), made up of sixty subunits called protomers. Although the capsid proteins share no sequence homology they share a common structural core in the shape of a triangular wedge consisting of an eight stranded antiparallel beta barrel or 'jelly-roll' (Hogle *et al.*, 1985). Length and

conformation of the loops connecting the beta barrel strands provide the greatest differences between the capsid proteins.

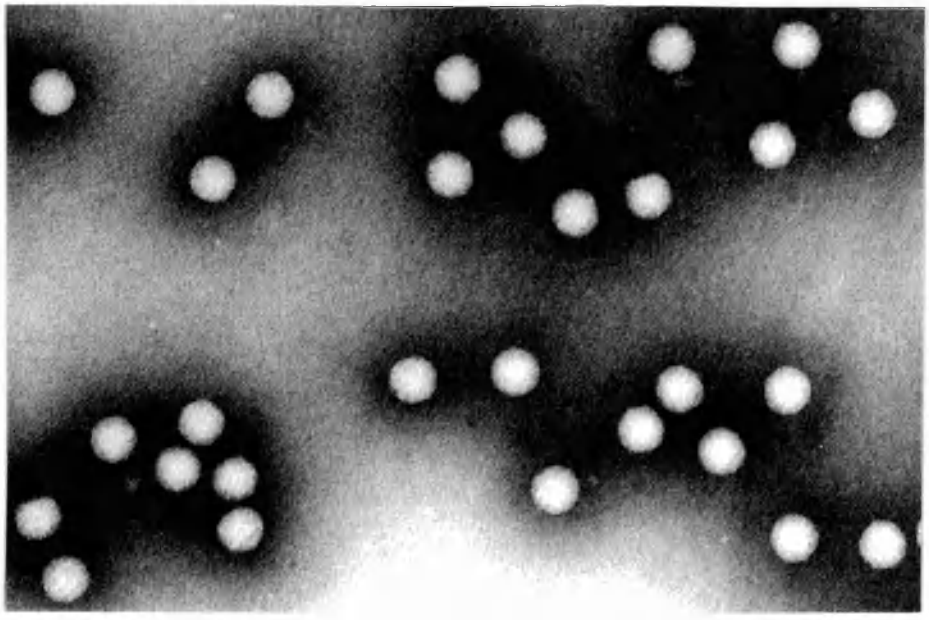


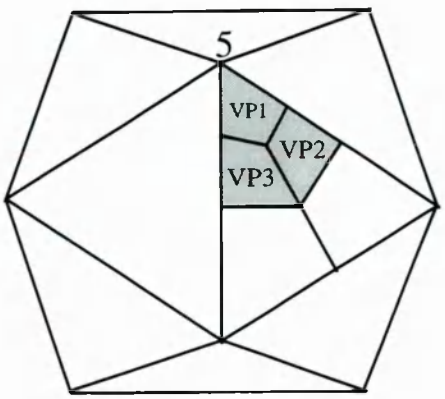
FIGURE 1.3 Transmission electron micrograph of poliovirus at a magnification of X 210,000. The specimen was prepared with negative staining using 4% aqueous sodium silicotungstate. The virus particles are icosahdral in shape and approximately 27nm in diameter. Electron microscopy by Dr D Hockley, NIBSC.

1.2.3 ASSEMBLY OF THE VIRUS PARTICLE

The viral genome is translated as a single polyprotein which is initially cleaved autocatalytically during translation by the protease 2A which hydrolyses the peptide bond at its own amino terminus to separate the capsid precursor (P1) and the non structural precursor (P2P3) (Toyoda *et al.*, 1986). P1 is initially cleaved to give peptides VP0, VP1 and VP3 which together form a protomer. Protomers then form pentamers which are able to form empty capsids (Palmenberg 1982). Later cleavage of VP0 to give VP4 and VP2 during RNA encapsidation (Palmenberg 1982) allows complete assembly of pentamers resulting in the icosahedral structure, fig 1.4 and an increase in stability. VP1, VP2 and VP3 are ordered together on the outer surface with VP4 internal which is myristoylated

(Chow *et al.*, 1987; Paul *et al.*, 1987). The myristoyl group is necessary for viral infectivity (Marc *et al.*, 1989; Kräusslich *et al.*, 1990). A lipid, possibly sphingosine is inserted into each VP1 molecule.

A



B

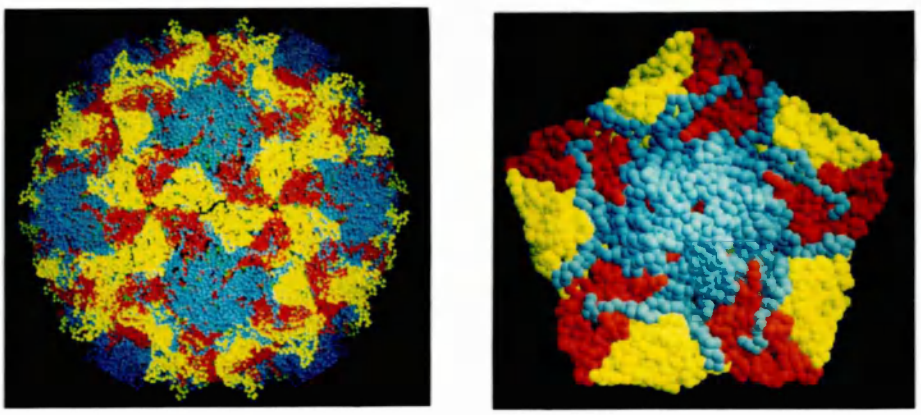


FIGURE 1.4 A. A geometric representation of the icosahedral structure of a poliovirus particle showing approximate positioning of capsid proteins VP1, VP2 and VP3, VP4 is internal. Fivefold symmetry at the pentameric axis is shown.
B. Space filling models showing the positions of the surface polyproteins. The figure on the left is of the entire virus particle and the figure on the right shows a single pentamer. VP1 is coloured blue, VP2 is yellow and VP3 is red. VP4 is internalised and therefore not shown.

1.2.4 INTERACTION WITH THE POLIOVIRUS RECEPTOR

A deep canyon surrounding the fivefold axis of each pentamer (indicated in fig 1.4) was suggested to act as a receptor attachment site (Rossmann *et al.*, 1985). This would place it relatively inaccessible to the host's antibodies avoiding structural changes arising from antigenic escape mutations. The equivalent canyon in rhinoviruses was confirmed as a receptor binding site by cryoelectron microscopy (Olsen *et al.*, 1993) but similar techniques have also detected a FAB (fragment antigen binding) particle of a strongly neutralising antibody bound into the canyon of a rhinovirus (Smith *et al.*, 1996). Binding to the receptor brings about a conformational change in the capsid that releases the RNA into the cytoplasm (de Sena & Mandel 1977). In contrast, a conformational change was found to occur in the FAB fragment on binding. Unlike the receptor, the antibody fragment was not found to reach the 'floor' of the canyon where amino acid sequence is highly conserved (Rossmann & Palmenberg 1988). The virus is therefore able to conserve the shape of the receptor binding site, required for subsequent entry of the virus into the cell, but still escape antibody recognition by mutating amino acids further up the wall of the canyon.

The human poliovirus receptor gene was cloned and its secondary structure deduced by examination of the amino acid sequence (Mendelsohn *et al.*, 1989). It was found to be an integral membrane protein and a member of the immunoglobulin superfamily. The long extracellular region at the amino terminus is thought to fold into three domains which are highly conserved across the family. The receptor for poliovirus was the third viral receptor to be identified as an immunoglobulin following CD4, the major receptor for HIV1, (Maddon *et al.*, 1986) and ICAM-1, the major rhinovirus receptor (Greve *et al.*, 1989; Staunton *et al.*, 1989). Recently the receptor for coxsackie B viruses and adenoviruses 2

and 5 was identified as yet another transmembrane immunoglobulin (Bergelson *et al.*, 1997).

On binding to its receptor poliovirus undergoes a large conformational change where VP4 is lost (de Sena & Mandel 1977) and the amino terminus of VP1 is externalised (Fricks & Hogle 1990). It is the exposed VP1 terminus which is responsible for liposome attachment, an action which is thought to facilitate the transport of viral elements across the cell membrane. The conformational change results in an increase in sensitivity of the virus to detergents and protease activity and a decrease in sedimentation rate of 160S to 135S (de Sena & Mandel 1977). The viral RNA is also sensitive to RNaseA in the presence of SDS but still encapsidated (Guttman & Baltimore 1977) and this altered (A) particle, eluted from cells is uninfectious (Holland 1962). This particle can however infect chinese hamster ovary cells or murine L cells which are normally non permissive (Currey *et al.*, 1996). The RNA is then released into the cell and becomes RNaseA sensitive without the presence of SDS, losing all capsid protein protection (Guttman & Baltimore 1977).

1.2.5 THE GENOME OF THE VIRUS

The poliovirus genome contains a relatively long NCR at the 5' end with an elaborate secondary structure and a shorter NCR at the 3' end followed by a poly-A tail. A single large open reading frame codes for the four structural proteins and the non structural proteins. A schematic of the genome is shown in fig 1.5. Many functional proteins that would appear to be precursors made up of two individual uncleaved proteins have also been identified as having specific activities during the life cycle of the virus.

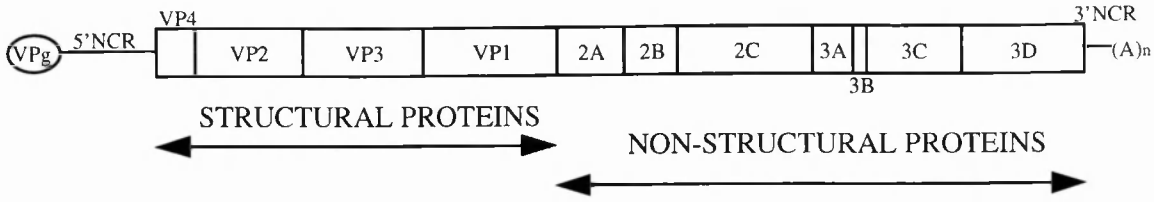


FIGURE 1.5 A schematic of the polioviral genome showing the positions of the individual protein genes, the noncoding regions and the 5'linked protein VPg.

While the polyprotein is translated, cleavage into the individual proteins is carried out by the proteases 2A, 3C and 3CD. 2A cleaves between tyrosine-glycine pairs and will be discussed later in greater detail. 3C cleaves between glutamine-glycine. The protease 3CD consists of sequences from both 3C and 3D and is involved in processing of the P1 precursor (Jore *et al.*, 1988). Recently 3CD has been implicated in the binding of a 38kDa cellular protein to the 3' end of the negative viral RNA strand during replication (Roehl *et al.*, 1997). The protein 3D is itself an RNA-dependent RNA polymerase and the 3D component of 3CD may alter the protease recognition site. 3C was also found to be able to bind to RNA (Andino *et al.*, 1993) and its role in replication is discussed later. The expression of 2B has also been shown to cause the disassembly of the golgi complex and the swelling of the endoplasmic reticulum (Sandoval & Carrasco 1997), it is the golgi complex that is the target for the anti-polioviral drug Ro (Ishitsuka *et al.*, 1982a; Ishitsuka *et al.*, 1982b). This supports the hypothesis that the replication of viral RNA is carried out on these membranes (Guinea & Carasco 1990). 2C is involved in RNA synthesis (Li & Baltimore 1988) and may have helicase activity but can be provided *in trans* and is also involved in determination of the virion structure (Li & Baltimore 1990). 3B, otherwise known as VPg, is bound to the 5' end of the viral genome but when it is present as 3AB it may act as a co-factor of 3D (Lama *et al.*, 1995). The protein 2BC was shown to block the exocytic pathway of *Saccharomyces cerevisiae* (Barco & Carrasco 1995), arresting cell

growth and stimulated the formation of small membranous vesicles in the cytoplasm of mammalian cells. Poliovirus RNA is thought to be translated and transcribed in these vesicles (Caliguri & Tamm 1969; Guinea & Carrasco 1990) and infected cells are stimulated to synthesise phospholipids which are used in vesicle formation (Mosser *et al.*, 1972). The function of 3A is still unknown although 3A and 2C have affinities for cell membranes (Takegami *et al.*, 1983; Bienz *et al.*, 1987).

1.3 THE 5' NONCODING REGION

1.3.1 SECONDARY STRUCTURE

The 5'NCR of polioviruses, approximately 740 nucleotides long, forms a very complex secondary structure as inferred by electron microscopy (Currey *et al.*, 1986), comparative sequence analysis and mutant-revertant analysis and validated by analysis with single and double strand specific reagents (Rivera *et al.*, 1988; Pilipenko *et al.*, 1989; Skinner *et al.*, 1989; Le & Zuker 1990). The region is divided into 6 distinct domains and a schematic is shown in figure 1.6 although more structural folding probably forms. The sequence of the 5'NCR is highly conserved within the poliovirus family and also across the enterovirus and rhinovirus genera. Comparison between members of the picornaviridae family largely reveal a structural conservation more than a sequence conservation (Jackson *et al.*, 1994) but all function in the same way. Sequence is however fairly well conserved between the three poliovirus serotypes.

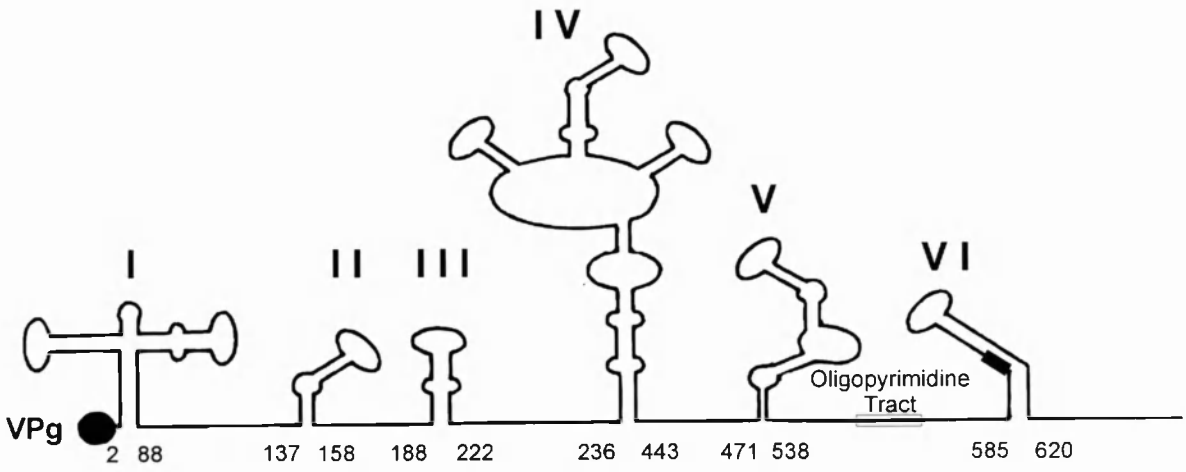


FIGURE 1.6 A representation of the secondary structure of the 5'NCR of polioviruses showing probable positions of domains I - VI, the VPg protein linked to the very 5' end, the oligopyrimidine tract and the initiating AUG codon in domain VI, depicted as solid rectangle (Pilipenko *et al.*, 1989; Skinner *et al.*, 1989; Le & Zuker 1990).

The 5'NCR of polioviruses contains both replication and translation signals and is an important regulatory element of the genome. Domain I forms a 'cloverleaf ' that is required in replication for synthesis of a positive sense strand of RNA although very little is known about the whole process. Domains II, IV, V and VI but not III together form the internal ribosome entry site or IRES which is responsible for directing ribosomes onto the RNA for internal initiation of translation.

1.3.2 DOMAIN V

Determinants of attenuation in all 3 Sabin strains of poliovirus are found in domain V of the 5'NCR (fig 1.7) and this domain forms part of the IRES. Attenuating mutations in the 5'NCR can also give rise to a cell specific temperature sensitive phenotype (Macadam *et al.*, 1992). They can all be rationalised as bringing about a change in the base pairing of the structure and disrupt the predicted secondary structure of the domain. This is verified in that revertants of these viruses that have become more virulent carry either direct back

mutations or second site mutations in the complementary strand allowing a new base pairing to retain overall shape. For example, in the type 1 virus, the base pair 480/525 (483/528 in type 3 numbering of fig 1.7) changes from A/U to G/U in Sabin 1 which attenuates the virus. Neurovirulent revertants were found to have a G to A change at 480 or a U to C change at 525 which would re-form the base pair as either AU or GC (Skinner *et al.*, 1989; Christodoulou *et al.*, 1990; Muzychenko *et al.*, 1991).

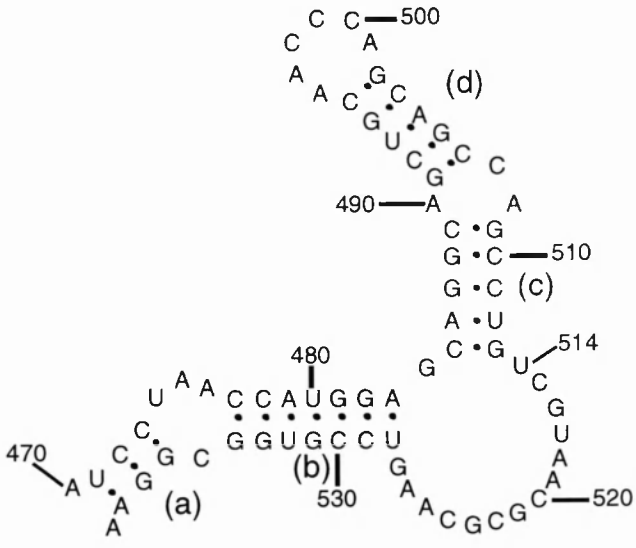


FIGURE 1.7 A representation of the predicted secondary structure of domain V of the 5'NCR of poliovirus using a type 3 Leon virus as example (Skinner et al., 1989). Stems are labelled for descriptive purposes.

Similarly a number of viruses were constructed with mutations around the stem formed by bases 466 - 478 and 533 - 540 of a type 3 domain V (Macadam *et al.*, 1992). By calculating the predicted free energy (ΔG) and the probable structures of the constructs it was possible to show that extent of disruption correlated with the extent of temperature sensitivity in BGM cells. For example simply swapping the base of 472/537 from C/G to G/C hardly affected the phenotype but replacement by the weaker base pair UA decreased stability of the structure and increased temperature sensitivity. Furthermore, different temperature sensitive phenotypes were found in different cell lines, implicating this structure in interaction with cell derived factors.

The secondary structure of the 5'NCR is therefore apparently very important to the virus. A correctly formed secondary structure would allow specific pockets and bulges to assemble so that interaction with both viral and cellular factors can occur for efficient replication and translation.

1.3.3. TERTIARY STRUCTURE

Little is known about the tertiary structure within the 5'NCR, the only evidence to date relies on computer predictions. Analysis of highly conserved sequence in enteroviruses and rhinoviruses has identified a possible pseudoknot (two loops of secondary structure which intertwine) at the 3' end of the 5'NCR (Le *et al.*, 1992). As part of the IRES the pseudoknot could be responsible for binding to cellular proteins in the initiation of translation and the pyrimidine rich stretch downstream from this could act in a similar way to the Shine-Dalgarno sequence of prokaryotic cells. A similar structure was later predicted for EMCV, FMDV, hepatitis A virus and Theiler's murine encephalomyelitis virus (Le *et al.*, 1993).

In addition to this another pseudoknot was also suggested to form between domains IV and V on observation of coupled mutations generally found in Sabin 2 (Muzychenko *et al.*, 1991). The changes at 398 and especially at 481 are both determinants of attenuation but are generally found to have reverted simultaneously in viruses excreted post immunisation. However no analogous coupled changes happen in the other Sabin strains giving little evidence to support this formation.

1.4 REPLICATION

The cloverleaf structure of domain I contains replication signals. Interaction with the viral protein 3CD and a cellular protein forms a membrane associated complex that is required for synthesis of positive sense RNA from a negative sense template. In contrast the cloverleaf structure is not required in the negative strand (Andino *et al.*, 1990; Andino *et al.*, 1993). The cloverleaf formed by the positive sense strand appears to be a requirement *in trans* for this complex to form before replication can occur and can of course be provided by the RNA of the infecting viral particle for the initial round of synthesis. The complex is not essential for negative strand synthesis.

There are a number of viral proteins involved in replication. Fundamentally 3D is a template and primer dependent RNA polymerase (Flanegan & Van Dyke 1979). 3CD binds near the 5' end of viral RNA in the replication complex (Andino *et al.*, 1990; Andino *et al.*, 1993) and 3CD/3C is also thought to be responsible for proteolytic modification of a cellular protein (Roehl *et al.*, 1997). Only the modified 38kD form of this protein appears to have a high affinity for binding to the 3'terminus of the viral negative RNA strand. From deletion mutants, residues 5-10 of this strand were found to be essential for viral replication (Roehl *et al.*, 1997). In addition the protease 3C is responsible for inhibiting transcription by cellular polymerase III, converting the active form of transcription factor IIIC to an inactive form (Clarke *et al.*, 1991). 2C has a nucleotide binding motif (Barton *et al.*, 1995) and 2B may be required *in cis* as 2BC (Johnson & Sarnow 1991) but the requirement in replication of these proteins is unknown. The membrane associated precursor 3AB may act to deliver 3B (VPg) (Giachetti & Semler 1991) but it has also been implicated as a co-factor of the polymerase 3D (Lama *et al.*, 1995).

1.5 TRANSLATION

Most cellular mRNAs have a capped extreme 5' end of m^7GpppNp structure which is utilised in cap-dependent protein synthesis. According to the scanning model (Kozak 1978, Kozak 1989) the initiation complex binds to the RNA at the cap structure, moves along and synthesis begins at the first AUG codon in the correct context on the RNA. In contrast, poliovirus RNA terminates at the 5' end with pUp (Hewlett *et al.*, 1976; Nomoto *et al.*, 1976). A small virally encoded protein, VPg, is attached to the 5' end and is cleaved off before translation begins. The virus instead uses a cap-independent process for protein synthesis where the secondary structure of the 5'NCR acts as an internal ribosome entry site (IRES) and protein synthesis begins at the 9th AUG codon. This alternative translation initiation allows polioviruses to shut off host cell cap-dependent translation (see below), freeing the host cell translation machinery to synthesise viral proteins.

1.5.1 THE INTERNAL RIBOSOME ENTRY SITE

Evidence for initiation of protein synthesis by internal ribosome entry stemmed from the construction of bicistronic constructs with 2 reporter genes separated by the 5'NCR of a picornavirus (Pelletier & Sonenberg 1988, Jang *et al.*, 1988). In infected cells the downstream cistron was found to be translated in greater amounts than the upstream cistron under the control of a capped RNA leader where as in the absence of the 5'NCR the downstream cistron was translated very inefficiently. This is consistent with the idea that few ribosomes are able to scan through the intercistronic region once translation of the first cistron has finished and that the secondary structure of the NCR would completely block scanning. Confirmation that internal initiation of translation does not require a free RNA 5' end was provided when a circular construct containing an EMCV IRES was prepared

(Chen & Sarnow 1995). Translation from the open reading frame placed after the IRES was achieved *in vitro* without linearisation.

The position of the IRES in the 5' NCR lies upstream of the initiating AUG codon which is found at position 743 in type 3 polioviruses and has been located by the construction of deletion mutants (Iizuka *et al.*, 1989; Pilipenko *et al.*, 1989; Nicholson *et al.*, 1991; Percy *et al.*, 1992; Haller *et al.*, 1993). The IRES spans a large part of the 5'NCR, incorporating domains II, IV, V and some of domain VI (fig 1.6). The conserved U-rich tract found between domains V and VI acts as a length specific spacer and is thought to allow the correct AUG codon to be brought into alignment. The ribosome therefore appears to be directed into place by the three dimensional structure formed by the 5'NCR. The unpaired bases found at the loop segments which are highly conserved are thought to interact either with the ribosome or other trans-acting proteins (Jackson *et al.*, 1994) or RNA in the tertiary structure. Defective IRES elements contained in bicistronic constructs can however be complemented *in trans* by the coexpression of the poliovirus genome (Percy *et al.*, 1992; Stone *et al.*, 1993). The complementation required a very high level of sequence conservation as only intact poliovirus IRES elements were able to successfully rescue a defective poliovirus IRES elements. This suggested interactions between RNA molecules was possible or that transfer of the assembled ribosome initiation complex to the defective IRES element can occur.

There are 3 distinct IRES and 5'NCR structures found among picornaviruses. The first is found in polioviruses, coxsackieviruses, rhinoviruses and enteroviruses and the second is found in cardioviruses and aphthoviruses. The third, found in hepatitis A virus is significantly different to exist in its own group (Jackson *et al.*, 1994). The IRESes have slightly different functional ribosome entry sites. The first group which includes

polioviruses direct the translational complex onto a non initiating AUG triplet which then moves to the authentic triplet, probably in a scanning like way (Pelletier & Sonenberg 1988). In contrast the second group of IRESes direct the complex directly to the authentic AUG triplet (Kaminski *et al.*, 1990). The hepatitis A IRES is twenty five times less efficient than an EMCV virus in tissue culture and this is thought to reflect a reduction in affinity for cellular factors and a possible reason why the virus grows so slowly *in vitro* (Brown *et al.*, 1994).

The cardiovirus and aphthovirus RNAs are very efficient templates of translation in rabbit reticulocyte lysates whereas poliovirus, enterovirus and rhinovirus RNAs are translated inefficiently and incorrectly due to requirement of cellular factors from nucleated cells. In artificially made constructs the IRES elements of EMCV can initiate translation from poliovirus RNA *in cis* with no detrimental effect on virus growth (Alexander *et al.*, 1994). Mutant forms of the EMCV IRES elements were also complemented *in trans* by coexpression of a positive sense intact EMCV IRES element with no evidence of recombination (Roberts & Belsham 1997). In addition the cloverleaf and the IRES seem to be functionally independent and non overlapping elements of the 5'NCR as constructs containing the cloverleaf of poliovirus and the IRES of EMCV are fully functional (Rohll *et al.*, 1994).

The initiation of translation of polioviral RNA requires trans-activating factors in addition to canonical initiation factors. This was apparent from *in vitro* translation studies using rabbit reticulocyte lysates as a translational model. Proteins synthesised were found to be unlike proteins produced *in vivo* and initiation was found to occur towards the 3' terminus. This phenomenon was rectified with the addition of a HeLa cell or Krebs-2 cell extract

which provided a cellular factor or factors missing or in limiting quantities in the lysates (Brown & Ehrenfeld 1979; Dorner *et al.*, 1984; Phillips & Emmet 1986). Addition of this factor was also found to be less active for RNA from an attenuated type 2 virus than from a neurovirulent type 2 virus, highlighting the importance of the change at residue 481 in attenuation and translation (Svitkin *et al.*, 1988).

1.5.2 INTERACTION OF CELLULAR PROTEINS WITH THE 5'NCR

There have been numerous accounts of proteins binding to various portions of the 5'NCR of polioviral RNA but the most recent work involves a 52kDa protein that bound to RNA at resides 559-624 and was found to be present in HeLa cells but not rabbit reticulocyte lysates (Meerovitch *et al.*, 1989). This protein was then identified by protein sequencing to be the La autoantigen (Meerovitch *et al.*, 1993) which is a human antigen recognised by antibodies in sufferers of systemic lupus erythematosus and Sjögren's syndrome (Tan 1989). It functions in transcription of RNA by polymerase III *in vitro* (Gottlieb & Steitz 1989) and contains an RNA binding site (Kenan *et al.*, 1991). A recombinant form of this protein was found to stimulate the translation of poliovirus RNA in the lysates and correct the aberrant initiation at the same time, although at levels far in excess of those present in the cell.

Investigations have found that a deletion of the C-terminal half of the La protein abolishes its effect on translation but still enabled binding to the 5'NCR of poliovirus. The protein is usually found in the nucleus but appears to move to the cytoplasm during poliovirus infection where it would be utilised. In addition, out of a number of proteins tested, only La, guanine nucleotide exchange factor (GEF) and eIF-2 were found to stimulate initiation of polioviral protein synthesis at the authentic site at the 5' end (Svitkin *et al.*, 1994). It has

also been reported that immunodepletion of La from HeLa cell lysates represses poliovirus translation although this effect could not be rescued by supplementation with La (Belsham & Sonenberg 1996).

A 57 kDa protein (p57) was also found to bind the 5'NCR of poliovirus. It was initially found by mutation of the conserved oligopyrimidine tract at the position 559-565 and the use of UV-cross linking with extracts from HeLa cells (Pestova *et al.*, 1991). Translation of poliovirus RNA was inhibited by competition with the p57 binding domain from the 5'NCR of EMCV (Pestova *et al.*, 1991) but p57 does not stimulate internal initiation of translation on human rhinovirus RNA by itself (Borman *et al.*, 1993). The same protein was also found to interact with the IRES of foot and mouth disease virus and was detected in HeLa cell extracts as well as in rabbit reticulocyte lysates (Luz & Beck 1991).

This 57kDa protein was found to be the pyrimidine tract-binding protein (PTB) (Hellen *et al.*, 1993) which is a cellular protein with at least two binding domains found to bind to various RNA or ssDNA sequences (Brunel *et al.*, 1996). Purified recombinant PTB bound specifically to a bulged hairpin of the IRES of EMCV more strongly than to a mutated version and immunodepletion of the protein from HeLa cell extracts inhibited translation of poliovirus RNA. PTB was also found to compete with human glyceraldehyde 3-phosphate dehydrogenase (GAPDH) for binding to the IRES of hepatitis A virus (Schultz *et al.*, 1996). Human GAPDII is thought to have an adverse effect on internal initiation of translation because it destabilises RNA secondary structure.

Although three binding sites have been identified in the 5'NCR of poliovirus, at 70-288, 443-539 and 630-730, (Hellen *et al.*, 1994) the function of PTB in translation is still not

clear. The attenuating mutation of A to G at 480 does not alter the UV cross linking of PTB to domain V and the 630-730 region can be deleted in poliovirus RNA without effect in tissue culture. It is suggested that PTB is responsible for bringing a number of regions of the 5'NCR together to promote interaction of the 40S ribosome subunit with the RNA in the 5'NCR (Pestova *et al.*, 1991) or it could be acting as a chaperone. But as with all protein binding assays, evidence that a protein will interact with RNA from UV cross linking does not necessarily mean it is functionally significant.

Recently La, along with other RNA binding proteins hnRNP A1 and PTB have been identified as able to prevent spurious intiations at aberrant translation sites in cap-dependent translation (Svitkin *et al.*, 1996). It is proposed that these proteins bind to the RNA and prevent the ribosomal complex from recognising other AUG codons. This could explain how translation systems which lack such factors allow incorrect translation of poliovirus RNA and gives La a role in both types of translation initiation.

1.5.3 INTERACTION OF CELLULAR TRANSLATION INITIATION FACTORS

From work using the EMCV IRES, cap-independent initiation of translation was found to require the same canonical initiation factors as cap-dependent initiation of translation (Pestova *et al.*, 1996a). This involved primer extension to confirm accurate initiation complex formation with and without the addition of purified components and initiation factors. From this, purified eIF2, eIF3, eIF4F, Met-tRNA$_i^{Met}$ (the initiating tRNA molecule carrying methionine), ATP, GTP and 40S ribosomal subunits were found to be required for the 48S complex to form on the IRES *in vitro*. eIF4B and PTB were not found to be essential but were able to enhance complex formation. This meant that neither mechanism

of initiating translation required other specific factors and implied that the IRES structure itself, not IRES-specific factors is able to direct internal entry of ribosomes.

In cap-dependent initiation eIF2F promotes binding of the initiating Met-tRNA$_i^{Met}$ to 40S subunits and this is stabilised by eIF3, forming a 43S complex. eIF3, along with other factors, is then responsible for promoting binding of the 43S complex to the RNA. The cap-binding complex eIF4F then assists the 43S complex to bind near the RNA 5' cap by interacting with eIF3. In cap-independent initiation the complex binds internally to the IRES.

eIF4F itself is made up of three subunits: eIF4A, a helicase; eIF4E, the cap-binding protein and eIF4G (formerly called p220, eIF-4γ, eIF-4Fγ). Cap-independent initiation on an EMCV IRES was found to require only the eIF4A and eIF4G subunits (Pestova *et al.*, 1996a). Infection with picornaviruses results in the inhibition of cap-dependent translation which correlates, for rhinoviruses, aphthoviruses and enteroviruses, with the cleavage of eIF4G by a viral protease to allow viral cap-independent translation to take precedence in the cell (Etchison *et al.*, 1982; Etchison & Fout 1985; Kräussllich *et al.*, 1987; Devaney *et al.*, 1988).

The functional domains of eIF4G were mapped by using purified cloned deletion mutants *in vitro*. Cap-independent initiation was found to require only the central third of the protein which was found to specifically bind to an EMCV IRES up-stream of the initiation codon (Pestova *et al.*, 1996b). eIF4E interacts toward the amino-terminus of eIF4G while eIF3 and eIF4A interact towards the carboxy-terminus. eIF4A however is able to bind to the IRES without the carboxy-terminus of eIF4G. It is eIF4A that unwinds RNA, probably downstream of where it binds, to allow entry of a single stranded region into the binding

site of the 40S ribosome subunit (Pause *et al.*, 1994a). This requires ATP (Ray *et al.*, 1985; Pestova *et al.*, 1996a).

1.5.4 CLEAVAGE OF eIF4G

The enteroviruses and rhinoviruses cleave eIF4G with the protease 2A whereas the aphthoviruses (FMDV) cleave with the unrelated L protease (Medina *et al.*, 1993). The cleavage sites are only 7 residues apart and have the same effect (Kirchweger *et al.*, 1994). Cleavage of eIF4G by 2A and L separates eIF4E, which is bound to eIF4G, from the rest of the cap binding complex, removing the ability of the complex to bind to the cap structure and cap-dependent translation is inhibited (Lamphear *et al.*, 1995) (see fig 1.8). In the picornaviridae family, only the enteroviruses, the rhinoviruses and the aphthovirus induce shut off of host cell translation by eIF4G cleavage. Hepatitis A virus and some echoviruses (orphanoviruses) do not stimulate an inhibition of host cell translation in infected cells (Coller *et al.*, 1990) and no eIF4G cleavage is detected (Coller *et al.*, 1991). The cardioviruses are thought to inhibit cell translation by simple competition of viral and cellular RNA but may also induce a change in ion concentration to assist this (Svitkin *et al.*, 1978; Mosenkis *et al.*, 1985; Duke *et al.*, 1992).

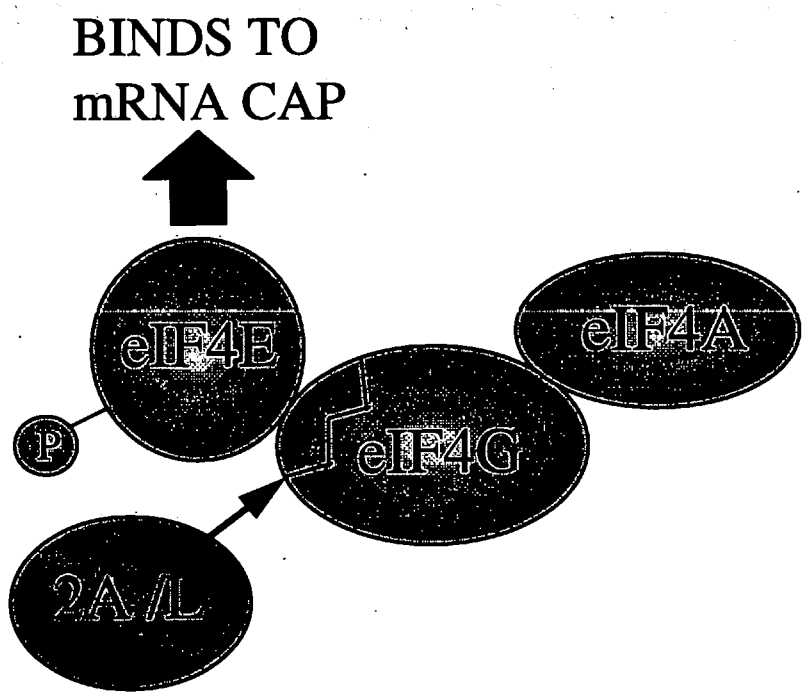


FIGURE 1.8 A simplified diagram to show the cleavage effect of 2A and L proteases on eIF4G of the initiation complex. eIF4G is lost from the complex which results in loss of cap-binding ability. Adapted from Belsham & Sonenberg (1996).

Initial evidence indicated that poliovirus protease 2A altered an unknown cellular protease which then cleaved eIF4G because the 'p220ase' cleavage activity seemed to be separate from 2A. Anti-2A serum failed to block eIF4G cleavage (Lloyd *et al.*, 1986) and eIF3 appeared to be required (Wyckoff *et al.*, 1990; Wyckoff *et al.*, 1992). This was overturned when coxsackievirus and rhinovirus recombinant 2A was found to be able to directly cleave the factor without the presence of eIF3 (Lamphear *et al.*, 1993). It is conceivable that poliovirus 2A acts in a different way but unlikely because of the high homology between the proteins and the difference could be due to poor purification of proteins in the initial experiments. eIF4G also has a binding site for eIF3 which could give confusing co-purification results.

Complete removal of the RNA coding for the leader protein in FMDV produces a viable virus but one which has a slower replication cycle, a delayed initiation of viral protein synthesis and is less efficient in host cell protein synthesis shut off (Piccone *et al.*, 1995). A virus of this construction has been found to be a possible attenuated vaccine candidate with promising results in tests with Hereford steers (Mason *et al.*, 1997). The FMDV leader protein, unlike the poliovirus 2A protein is found only to cleave eIF4G and has no role in polyprotein processing. Therefore shut off of host cell protein synthesis is not absolutely essential for viral proliferation. In contrast the removal of 2A and the insertion of an EMCV IRES in its place renders a virus non viable (Molla *et al.*, 1993b). This means that 2A has a role other than in autocatalytic cleavage.

1.5.5 ALTERNATIVE MECHANISMS FOR INHIBITION OF CELLULAR TRANSLATION

Cleavage of eIF4G alone only inhibits 70% of host cell protein synthesis whereas infection with poliovirus results in complete inhibition (Bonneau & Sonenberg 1987a; Pérez & Carasco 1992). Other possible mechanisms may therefore occur. Recently two small proteins have been found that bind to the cap binding protein, eIF4E, once one of them has been dephosphorylated (see fig 1.9). This results in disassociation of eIF4E from the cap binding complex and inhibition of cap-dependent translation (Pause *et al.*, 1994a). Dephosphorylation of this protein, 4E-BP1, has been found in poliovirus infected cells and interaction of it with eIF4E does not seem to affect cap-independent translation of poliovirus (Pause *et al.*, 1994b). This process was also found to occur in cells infected with EMCV (Gingras *et al.*, 1996). EMCV causes host cell translation to be shut off but it occurs later in the infection cycle than that found in poliovirus infection and no virus induced cleavage of eIF4G occurs (Mosenkis *et al.*, 1985).

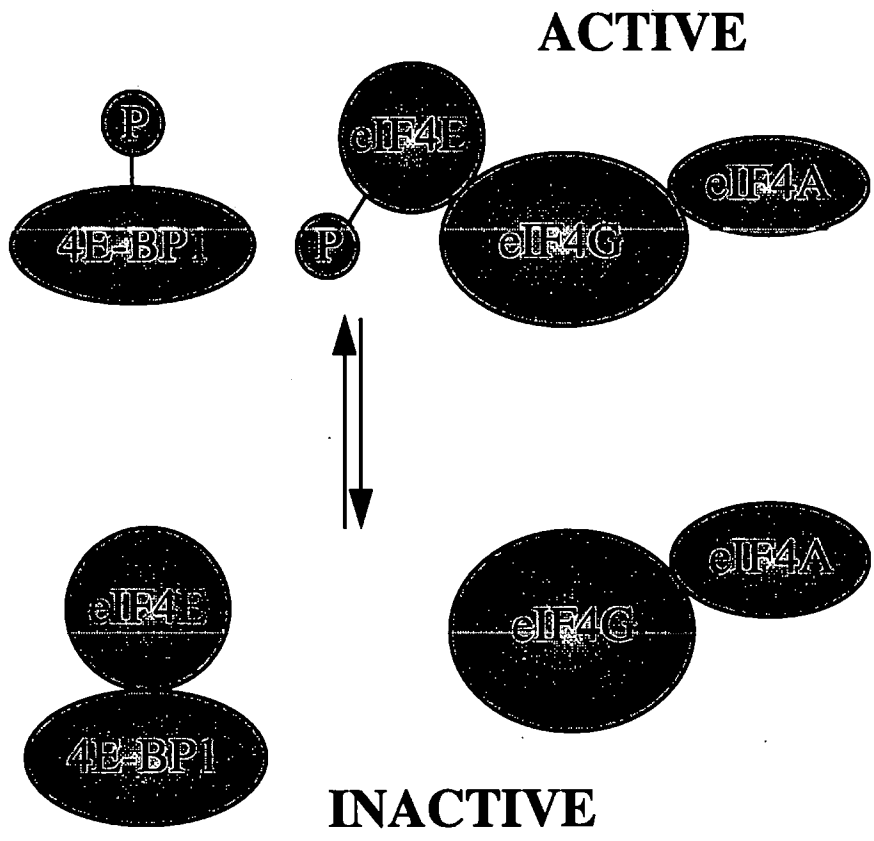


FIGURE 1.9 Diagram to show an alternative method of inactivation of the cap-binding ability of the cap-binding complex. Dephosphorylation of 4E-BP1 allows it to bind to the cap-binding protein eIF4E which dissociates from the complex. Adapted from Belsham & Sonenberg (1996).

An alternative mechanism of inhibition of cellular translation may be phosphorylation of the initiation factor eIF2α, a subunit of eIF2, which has been shown to occur in poliovirus infected cells (Black *et al.*, 1989; O'Neill & Racaniello 1989). Phosphorylation of this factor prevents its recycling and has an inhibitory effect on cap-dependent translation but may also affect cap-independent translation. eIF2 acts as a substrate for double-stranded-RNA-activated protein kinase (p68). It is thought that double stranded RNA synthesis during poliovirus infection induces high levels of p68 autophosphorylation although levels of p68 decrease. Subsequent phosphorylation of eIF2α results in a reduction in protein synthesis which then causes levels of p68 to be reduced. This elaborate system is thought

to exist because unlike viruses such as influenza, adenoviruses and vaccinia virus, poliovirus causes no direct inhibition of p68 (Black *et al.*, 1989).

1.5.6 INTERNAL INITIATION OF TRANSLATION

Picornaviruses are not the only RNA frameworks that can utilise internal initiation of translation. Other viruses and some cellular RNAs also have this ability. The human hepatitis C virus, a flavivirus, has a relatively long 5'NCR with elaborate predicted secondary structure (Brown *et al.*, 1992) which can direct internal initiation of protein synthesis in dicistronic constructs (Tsukiyama-Kohara *et al.*, 1992; Wang *et al.*, 1993; Fukushi *et al.*, 1994; Rijnbrand *et al.*, 1995; Reynolds *et al.*, 1995; Honda *et al.*, 1996). There is currently conflicting evidence as to minimum requirements of RNA needed for accurate initiation, possibly due to some interference from the RNA of reporter sequences in the dicistronic constructs. Unlike picornaviruses part of the coding region seems to be necessary (Reynolds *et al.*, 1995; Honda *et al.*, 1996) and internal initiation occurs at the initiating AUG codon like the cardiovirus and aphthoviruses and undergoes no scanning at all (Reynolds *et al.*, 1996). Indeed viruses classified as pestiviruses, a genus of the flavivirus family, all appear to have an IRES and utilise internal entry of ribosomes in translation initiation (Poole *et al.*, 1995).

Infectious bronchitis virus, a coronavirus, has a genome which is tricistronic with three open reading frames. Protein synthesis from the third is mediated by an internal initiation cap-independent mechanism (Liu & Inglis 1992). Also late mRNAs from adenoviruses have a common 200 base 5'NCR which can enhance translation of most mRNAs when it is added to their 5' ends. Adenovirus induced inhibition of host cell cap-dependent translation does not involve cleavage of eIF4G but inactivation by dephosphorylation of the cap-binding protein eIF4E (Huang & Schneider 1991). Recently the human leukeamia

virus was also found to contain an IRES-like sequence with evidence of a polypyrimidine region (Attal *et al.*, 1996). Retroviruses also contain other, shorter regions of RNA secondary structure to which viral proteins bind in a translation regulatory manner *e.g.* Rev protein of HIV-1 binds to the Rev response element (RRE) and is essential for expression of gag and pol proteins (reviewed in Cullen 1992).

The first mammalian RNA found to utilise internal initiation of translation was identified by Macejak & Sarnow (1991). Synthesis of the immunoglobulin heavy-chain binding protein (BiP) was detected during polioviral infection after cap-dependent translation had been inhibited and the 5' leader of the mRNA was found to be able to direct internal ribosome entry. Following this discovery, the homeotic gene antennapedia (Antp) mRNA of *Drosophila melanogaster* (Oh *et al.*, 1992), mRNAs in cell extracts from *Saccharomyces cerevisiae* (Iizuka *et al.*, 1994), human fibroblast growth factor 2 (FGF-2) mRNA (Vagner *et al.*, 1995) and eIF4G mRNA (Gan & Rhoads 1996) have all been discovered to initiate internal entry of ribosomes by RNA structures. In fact BiP, Antp and FGF-2 were all found to have a common RNA structural motif from computer folding analysis (Le & Maizel Jr. 1997).

The rationale of the ability of eIF4G mRNA being able to utilise cap-independent translation was proposed by Gan & Rhoads (1996). As the protein itself is involved in cap-dependent translation, any reduction in its concentration would reduce efficiency of this type of translation. eIF4G could then be synthesised using the alternative method in an autoregulatory mechanism. In addition, BiP shows some sequence homology with heat shock proteins and also increases in cells during stress as heat shock proteins do. It is

conceivable therefore that eIF4G concentration may also increase as a result of stress so that synthesis of specific proteins is not a limiting factor at such times.

Polioviruses would therefore appear to utilise a translation mechanism that eukaryotic cells are already able to employ when required. It is known that during mitosis the rate of protein synthesis is reduced to 25% of the rate during interphase and that the interaction of the cap-binding complex is reduced in mitotic cell extracts (Bonneau & Sonenberg 1987b). This mitotic block has no effect on poliovirus translation and phosphorylation of the cap-binding protein is implicated as a switch of this process. Furthermore phosphorylation of the cap-binding protein decreases when cells are heat shocked and addition of the protein during this state stimulates synthesis of non heat shock proteins.

1.6 THE VACCINES

In response to public feeling that the problem of poliomyelitis should be addressed, two different vaccines were formulated. Generally administered as trivalent preparations to confer immunity to all three serotypes, their success has been immense. The Salk inactivated poliovirus vaccine (IPV), administered intramuscularly, was first introduced in 1954 and the Sabin oral poliovirus vaccine (OPV) in 1960. These are still in use today.

Jonas Salk prepared the IPV from virus grown in monkey kidney cell cultures and inactivated with formalin. Field trails in Canada, Finland and the USA showed it to be safe and effective but problems arose immediately after its release. Inadequately inactivated vaccine was linked to the development of paralytic poliomyelitis and subsequent death of vaccinees in the 'Cutter incident' (Nathanson & Langmuir 1963). The first case of vaccine related poliomyelitis was admitted to hospital literally days after the vaccine was released

and the outbreak was traced to faults in production which were rectified. Since this incident there have never been any other IPV related cases and the vaccine is still regularly used.

Albert Sabin in contrast developed a vaccine that contained live virus and was administered orally, the route of natural infection. He prepared his vaccine by passaging virulent precursors of type 1 and type 3 viruses and a naturally attenuated type 2 virus isolated from wild type infections through a series of tissue cultures (Sabin & Boulger 1973). The virus therefore became adapted to conditions enforced through the tissue cultures which were different to those found in natural human infection. Sabin vaccine virus strains were thus developed that were selected for their low monkey neurovirulence and their ability to replicate in the human gut but retained antigenic determinants that could bring about an immune response.

It is necessary to batch test the Sabin vaccine to ensure that the virus has not altered in neurovirulence during production and a standard monkey neurovirulence test was introduced for this (WHO 1983). No other test or assay has been found that accurately provides a measure of the virulence. A number of vaccine related cases of poliomyelitis have however occurred. It is estimated that the incidence of occurrence is calculated at 1 per 5.3×10^5 for first time vaccinees and 1 per 2×10^6 overall (Nkowane *et al.*, 1987).

Inactivated and live vaccines each have their own advantages that influence their usage in different countries. The OPV avoids the use of needles which need to be sterile and can be traumatic for many adults and children and induces higher levels of secretory immunoglobulin A in the gut which is a similar response to wild type infection. The

mucosal immunity is however short lived and in contrast the IPV acts primarily to induce circulatory antibodies in the blood stream. The OPV brings about problems in its transportation as a cold chain has to be maintained from manufacture to immunisation. The use of deuterium oxide and magnesium chloride were tested to prevent heat inactivation but as many other vaccines are now included in the cold chain this is no longer a problem.

In addition there remains a residual pool of virus in the environment as long as the live vaccine is administered as oral vaccinees can excrete virus up to 73 days post immunisation (Minor *et al.*, 1986). Excreted type 3 virus was found to have reverted to a relatively high level of neurovirulence within 48 hours (Evans *et al.*, 1985). In contrast, where IPV is administered of course no excretion occurs. There is some evidence however that vaccination with IPV followed by OPV results in a greater frequency of faecal shedding than vaccination with OPV followed by OPV (Abraham *et al.*, 1993). This may prove a problem when the decision to stop vaccination world-wide is taken, the use of the OPV may be prohibited and only IPV will be used.

The understanding of the attenuation of polioviruses could only begin to be divulged when the appropriate molecular biology methods were developed. It may therefore be surprising that the Sabin vaccines have been in use for nearly four decades and attenuation is still not completely understood. Sequence comparison shows that there are many base changes produced during derivation of vaccine strains but the number of major attenuating determinants was found to be quite small, as summarised in fig 1.10 below. This fine balance reinforces the need for continued rigorous batch testing.

1.6.1 ATTENUATION

An enormous amount of research has been carried out to determine the important attenuating mutations. Much of the work has exploited the infectivity of cDNA clones (Racaniello & Baltimore 1981b) and RNA transcripts derived from them using a T7 promotor within the plasmid clone (Van der Werf *et al.*, 1986). 'Cut and paste' techniques between virulent precursors, virulent revertants for type 2, and attenuated viruses initially identified where attenuating regions of the genome were. Cloning techniques and site directed mutagenesis have given information of individual mutations and the detection of revertants has given further insight into the way in which the mutations affect the virus and an understanding of the stability of the vaccine.

The type 1 component of the vaccine is very rarely implicated in vaccine related poliomyelitis and this is probably due to the fact that it has 56 nucleotide differences when compared to the neurovirulent Mahoney strain. A relatively large number of mutations may be required for a revertant virus to arise and the vaccine strain is therefore relatively stable (Omata *et al.*, 1986). As a result of this type 1 attenuation is not completely resolved.

Originally three determinants of attenuation in Sabin 1 were identified: a change at 480 in the 5'NCR; a change in the 3D polymerase gene and the loss of a base at the start of the poly-A tail (Christodoulou *et al.*, 1990). This work used temperature sensitivity as a measure of neurovirulence but using a monkey neurovirulence test the change at 480 was identified as making an important contribution to attenuation with only a minor effect on temperature sensitivity (Kawamura *et al.*, 1989). This change of an A to G is thought to weaken the base pair of AU at 480/525 and destabilise the domain V structure. Using a

non transgenic mouse model and a mouse adapted version of Mahoney the change in 3D was also found to contribute albeit to a lesser extent than the 480 change (Tardy-Panit *et al.*, 1993). A further study using transgenic mice highlighted a number of changes in VP1, VP3 and VP4 that were thought to interfere with receptor binding and other associated events. This study also claimed that the changes in 3D and the 3'NCR were determinants of temperature sensitivity and not neurovirulence (Bouchard *et al.*, 1995). This was further substantiated in McGoldrick *et al.*, (1995) in which the mutations at 480 and in 3D were investigated in monkey neurovirulence tests and found not to be strong attenuating determinants.

The issue is somewhat complicated by the different ways of detecting neurovirulence although the monkey neurovirulence test must be decreed the most reliable as a measure of human neurovirulence. There have been no vaccine related epidemics since using this test. Temperature sensitivity merely measures an *in vitro* observation linked in some cases to neurovirulence. Attenuation of type 1 therefore appears to be due to a number of contributory changes with no single very strong determinant.

Attenuation for the remaining virus types is better understood. Type 2 polioviruses have two major determinants of attenuation. A change of G to A at position 481 in the 5'NCR which is thought to disrupt RNA secondary structure in domain V and an amino acid change at position 143 in VP1 caused by a change in residue 2908 which possibly interferes with receptor mediated events (Ren *et al.*, 1991; Macadam *et al.*, 1993). The change to A at 481 is thought to make a base pair with the U at 511. This is usually unpaired and would put extra strain on the domain V structure.

For the type 3 virus a change of C to U at position 472 in the 5'NCR is a determinant of attenuation (Westrop *et al.*, 1989) and is responsible for a translation deficiency (Svitkin *et al.*, 1990). A weakening of the base pair from CG to UG in one of the domain V stems is thought to destabilise the structure. Another mutation at residue 2034 which changes amino acid 91 in VP3 from serine to phenylalanine was also found to attenuate the virus (Westrop *et al.*, 1989). This mutation is thought to inhibit the assembly of protomers into pentamers (Macadam *et al.*, 1991a). A third mutation at position 2493 changes amino acid 6 in VP1 from an isoleucine to a threonine in the vaccine strain (Weeks-Levy *et al.*, 1991) which probably affects either the assembly of the capsid or uncoating of the RNA.

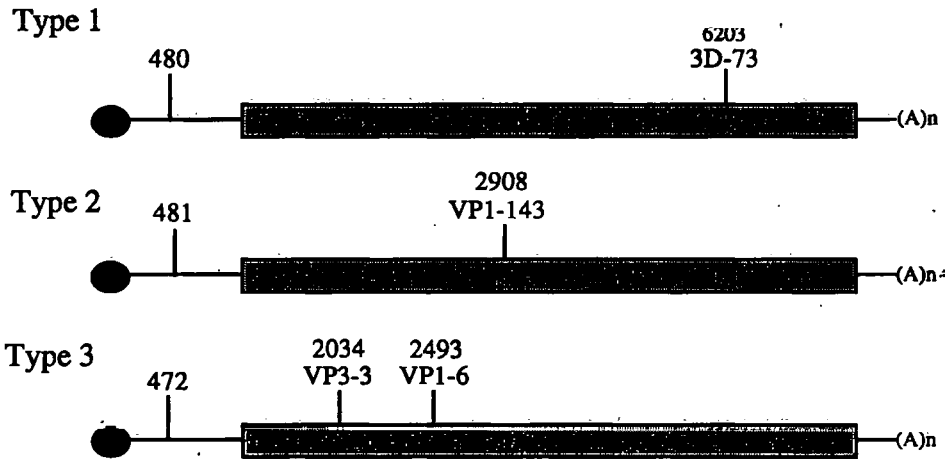


FIGURE 1.10 A schematic diagram of the 3 poliovirus serotypes indicating positions of some of the base changes associated with attenuation.

Evolution of the vaccine strains in the human gut involves more than point mutations. As the vaccine is administered as a trivalent preparation a single cell could be infected with more than one virus at any one time. This can give rise to intertypic recombinants whose functions are unknown but have been suggested as an evasion of the immune system (Minor 1993). Recombination was detected between type 2 and type 3 Sabin strains with cross-over occurring in the genome coding for non-structural proteins (Minor *et al.*, 1986).

On following the excretion of type 3 virus from a primary vaccinee a typical profile was obtained: type 3/type 1 and type 3/type 2 recombinants were detected at days 10-11 with a cross-over in the gene for 2C and type 3/type 2/type 3 recombinants were then detected at day 28 with a further cross-over in the gene for 3D (Cammack *et al.*, 1988).

1.7 THE VIRAL PROTEIN 2A

Polioviruses code for 2 proteases, 2A and 3C which process the polypeptide into individual proteins during the viral replication cycle. 2A is a small serine like protease, with a cysteine in place of the serine in the active site, of just 149 amino acids in length, cleaving at tyrosine-glycine peptide bonds. Only two of the ten possible tyrosine-glycine bonds are cleaved which indicates interaction with unknown factors, possibly structural cues. Mutational analysis has highlighted the threonine residue two bases away from the cleavage site as being important for cleavage (Lee & Wimmer 1988).

A putative model (fig 1.11) of the α-carbon chain was deduced from comparison with small bacterial serine proteases (Bazan & Fletterick 1988) with the putative catalytic triad being residues 20, 38 and 109. Residues 88, 89 and 124 are possibly involved in substrate binding (Hellen *et al.*, 1991; Yu & Lloyd 1991). A zinc binding site was detected on the 2A of human rhinoviruses and zinc depletion was shown to inactivate the virus (Sommergruber *et al.*, 1994). Zinc binding sites are found on other serine proteinases. Similarly the proteolytic activity of poliovirus and rhinovirus 2A was inhibited by elastase-specific inhibitors elastinal and methoxysuccinyl-Ala-Ala-Pro-Val-chloromethylketone (Molla *et al.*, 1993a). The substrate binding pocket of 2A is presumably comparable to the elastase binding pocket. The amino acid sequence of the protease is fairly well conserved with nearly 90% identity within polioviruses.

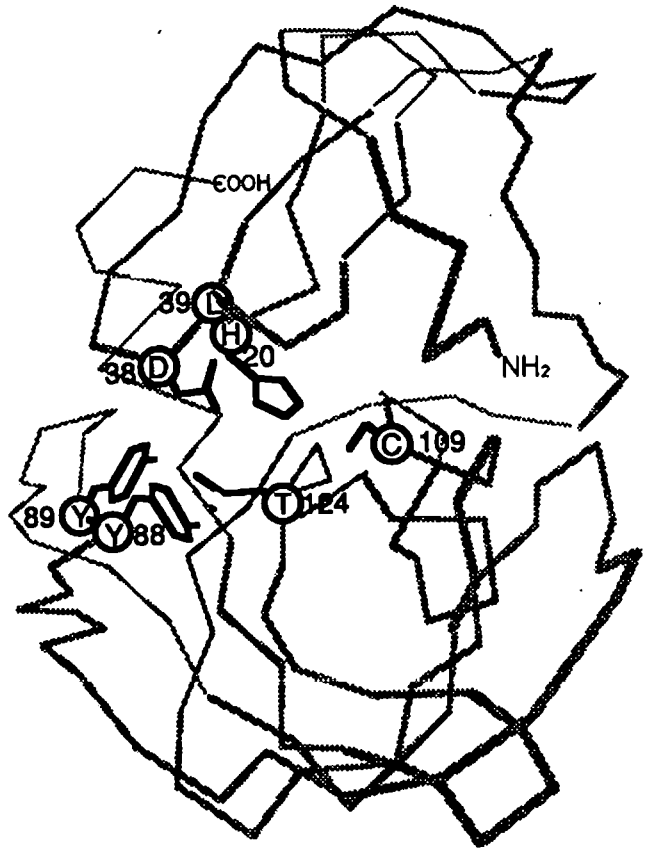


FIGURE 1.11 Predicted structure of the 2A poliovirus protease based on structure of small bacterial serine proteases (Bazan & Fletterick 1988; Yu & Lloyd 1991). Residues 20,38 and 109 represent the putative catalytic triad and residues 39, 88, 89 and 124 represent highly conserved amino acids.

2A cleaves itself autocatalytically from the polyprotein at the interface of VP1 and its own amino terminus. This occurs before complete synthesis of the polyprotein, separating the capsid precursor and the non capsid precursor and is the initial proteolytic event. 2A is also responsible for an alternative cleavage in the polyprotein during processing in some viruses. The protease 3C routinely cleaves to give 3C and 3D but in this case 2A cleaves further into the 3D part of the protein to give 3C' and 3D' (Lee & Wimmer 1988). In addition, acting again as a protease, 2A is responsible for shut off of host cell translation as discussed above where it directly cleaves the eukaryotic initiation complex subunit eIF4G.

1.7.1 2A AND THE 5'NCR

For the majority of viruses reversion of the ts phenotype that results from disruption to domain V in the 5'NCR involves re-forming the structure as previously described. This can be visualised by analysis of protein synthesis where a base pair disruption brings about a reduction in efficiency of synthesis in BGM cell culture at elevated temperatures. When the base pair is re-formed efficiency of synthesis is restored. This is demonstrated in fig 1.12 where base pair strength at 472/537 correlates with efficiency of protein synthesis at 38.5°C (personal communication, A. J. Macadam).

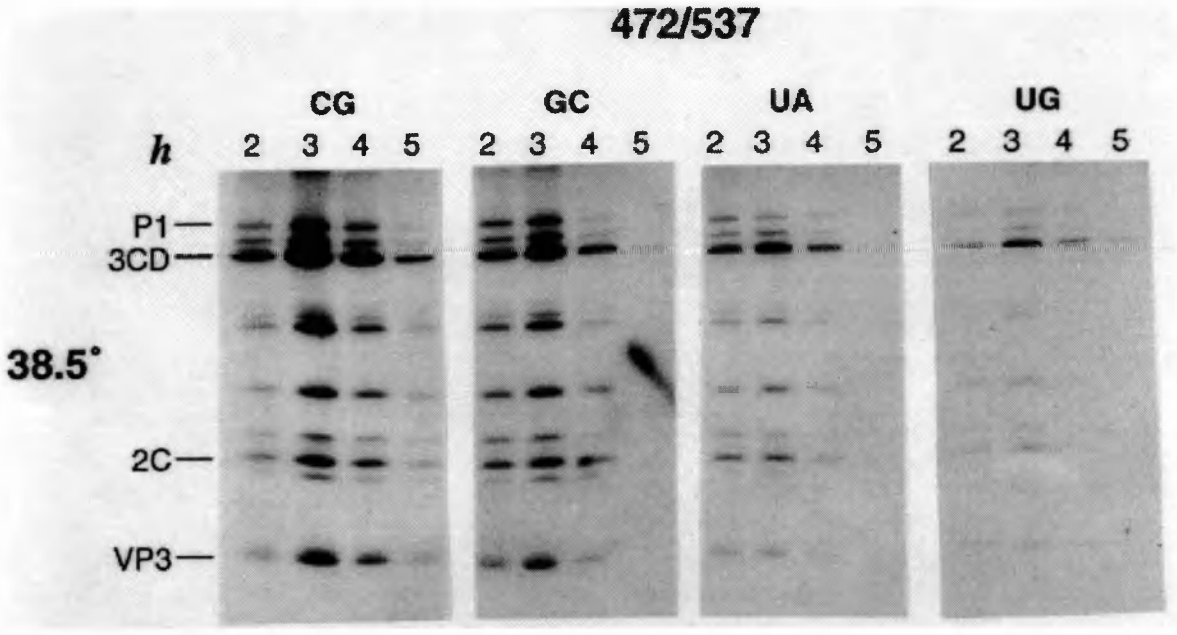


FIGURE 1.12 Protein synthesis at 38.5°C in BGM cells infected with viruses that differ only at the base pair 472/537 as indicated. Infected cells were pulsed with S^{35} Methionine for 30 min at 2, 3, 4 and 5 hours post infection. Cell lysates are then run on acrylamide gels.

However non ts viruses were selected which retained their domain V disruption but were no longer ts (Macadam *et al.*, 1994). Parental viruses were either Sabin 2 or a type 3/type 2 Leon/Lansing construct that had a range of individual domain V mutations created by site-directed mutagenesis. The revertant viruses were found to have compensating mutations in the coding region for 2A that resulted in amino acid substitutions. The compensating mutations were found to be able to restore protein synthesis in BGM cell culture

comparable to that resulting from reversion in domain V at a time when host cell translation was already inhibited (see fig 1.13). It appeared that the protein 2A was responsible for enhancing protein synthesis in some way.

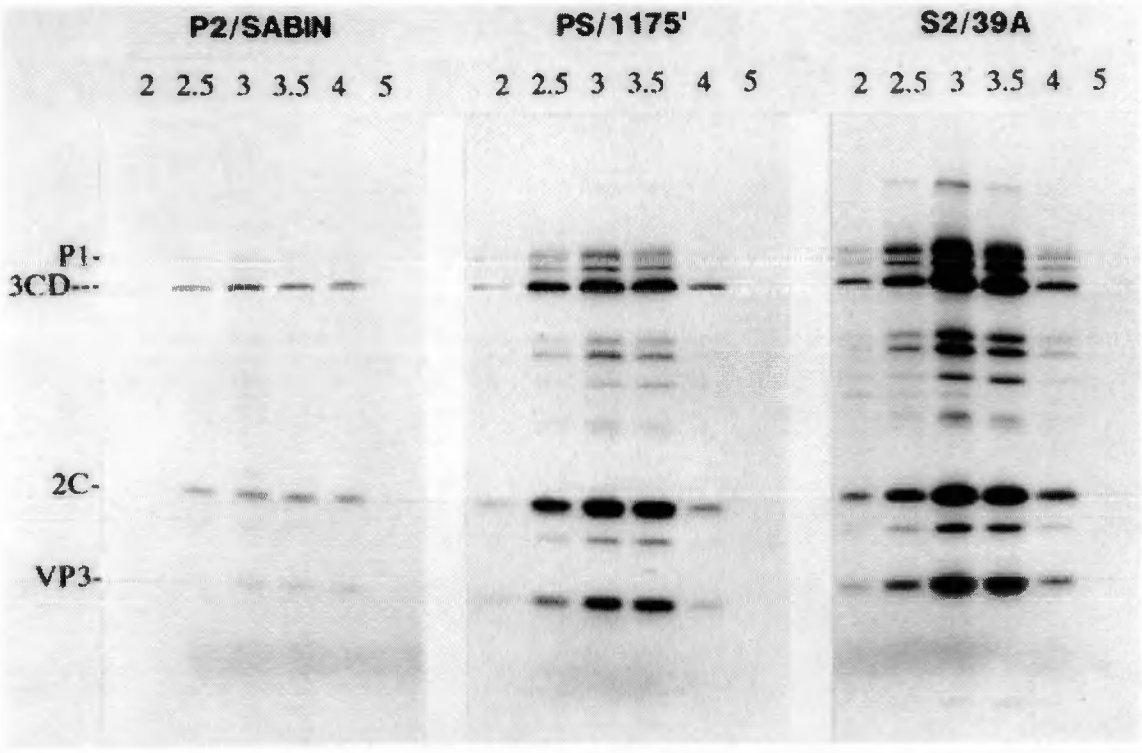


FIGURE 1.13 Protein synthesis at 38.5°C in BGM cells. Infected cells were pulsed with S^{35} labelled methionine for 30 min at 2, 2.5, 3, 3.5, 4 and 5 hr post infection. Viruses were Sabin 2 and derivatives. PS/1175' differed only in the 5' end and S2/39A differed only in 2A. Positions of viral polypeptides are indicated. Taken from Macadam *et al.*, (1994).

The levels of viral RNA in infected cells were comparable for all three viruses indicating that neither the disruption to domain V or the enhancing effect of 2A affected replication rates. The affect of 2A is therefore suggested to be on translation.

Earlier, Hambidge & Sarnow (1992) had suggested that 2A acted as a translational transactivator of poliovirus translation. Using bicistronic constructs of two reporter genes separated by the poliovirus IRES they detected enhanced translation of the second cistron in poliovirus infected cells even before cap-dependent translation had been inhibited indicating that cap-independent translation does not fundamentally require the cleavage of

eIF4G. Following this, translation from IRES driven RNAs was examined with expression of various parts of the poliovirus genome. Enhancement of translation of these RNAs was found to occur in cells in which 2A was expressed. As IRES elements can direct cap-independent translation of a reporter gene in rabbit reticulocyte lysates without the presence of polioviral proteins it therefore appears that 2A acts as an enhancer of cap-independent translation rather than an indispensable requirement.

For 2A to affect translation it would appear to have to interact with either the RNA of the 5'NCR to help form the tertiary structure or with the cellular proteins involved with translation. No evidence for 2A binding to the 5'NCR has yet been found. 2A was also shown to have a role in replication (Molla *et al.*, 1993b).

1.8 INTRODUCTION TO PROJECT

As discussed above, the Sabin vaccine strains have major attenuating mutations in domain V of the 5'NCR and are known to revert here during their passage through the gut of a vaccinee. This poses a problem of possible vaccine related poliomyelitis in primary vaccinees or their contacts. A more stable vaccine that does not revert would therefore be favourable. However weakening the 5'NCR of the vaccine strain in a more genetically stable way may not prevent reversion to virulence as mutations in the 2A protein were found to be able to compensate *in vitro* for the domain V mutations (Macadam *et al.*, 1994).

This thesis will describe an investigation into these compensating 2A changes initially discovered whilst studying domain V reversion. It will show that there are many positions and substitutions that can be found and that they all compensate a ts phenotype to different

extents. The cell specificity of 2A changes will be discussed along with the use of this in selection of domain V revertants. Important areas of domain V will be outlined from this selection and analysed using mutagenesis experiments. Results from these experiments will give information about functionally important areas of the 5'NCR domain V.

CHAPTER TWO

MATERIALS AND METHODS

2.1 SOLUTIONS AND MEDIA

2.1.2 PURIFICATION OF DNA

PHENOL MIX	50g phenol shaken with 50ml 10mM tris-HCl pH 8.6 0.1g dihyroxyquinoline
PHENOL/CHLOROFORM MIX	50ml phenol mix with tris buffer 50ml chloroform
TE BUFFER	10mM tris-HCl pH 7.5 1mM EDTA
SOLUTION 1	50mM glucose 25mM tris-HCl pH8 10mM EDTA
SOLUTION 2	0.2M sodium hydroxide 1% SDS
SOLUTION 3	3M potassium acetate pH 4.8
PREP-A-GENE BINDING BUFFER	6M sodium perchlorate 50mM tris-HCl 1mM EDTA pH 7.4
PREP-A-GENE WASH BUFFER	20mM tris 2mM EDTA 0.4M sodium chloride 50% ethanol pH 7.5

2.1.3 PURIFICATION OF RNA

15% SUCROSE SOLUTION	15% RNase free sucrose 33% normal saline 10mM tris-HCl pH 7.4
45% SUCROSE SOLUTION	45% RNase free sucrose 33% normal saline 10mM tris-HCl pH 7.4
SDS/SODIUM ACETATE	0.1% SDS 0.15M sodium acetate pH 6.0

2.1.4 SEQUENCING AND ELECTROPHORESIS OF DNA AND RNA

SEQUENCING GEL	6% acrylamide 42% urea 0.3% methylenebisacrylamide 0.12% ammonium persulphate 1 X TBE to set: 1µl TEMED ml^{-1} gel mix
WACKER'S SOLUTION	0.003% methacryloxy-propyltrimethoxysilane (Sigma) 0.003% glacial acetic acid in ethanol
ANNEALING BUFFER	280mM tris-HCl pH7.8 100mM magnesium chloride 350mM sodium chloride
LABELLING MIX FOR DNA SEQUENCING	2µM dGTP 2µM dCTP 2µM dTTP
LABELLING MIX FOR RNA SEQUENCING	2µM dITP 2µM dCTP 2µM dTTP
TERMINATION MIXES FOR RNA SEQUENCING	0.5mM dATP 1mM dITP 0.5mM dCTP 0.5mM dTTP FOR 'A' MIX: 0.25mM ddATP FOR 'I' MIX: 0.5mM ddITP FOR 'C' MIX: 0.25mM ddCTP FOR 'T'MIX: 0.25mM ddTTP
TERMINATION MIXES FOR T7 DNA SEQUENCING	127µM dATP 127µM dGTP 127µM dCTP 127µM dTTP 8.5mM magnesium chloride 38mM tris-HCl pH 7.5 42mM sodium chloride FOR 'A' MIX: 15µM ddATP FOR 'G' MIX: 15µM ddGTP FOR 'C' MIX: 15µM ddCTP FOR 'T' MIX: 15µM ddTTP

TERMINATION MIXES FOR DNA CYCLE SEQUENCING	15µM dATP 15µM dGTP 15µM dCTP 15µM dTTP FOR 'A' MIX: 600µM ddATP FOR 'G' MIX: 30µM ddGTP FOR 'C' MIX: 450µM ddCTP FOR 'T' MIX: 1200µM ddTTP
FORMAMIDE DYE MIX	20ml formamide (saturated with amberlite) 0.8ml 0.5M EDTA 6pg xylene cyanol 6pg bromophenol blue
10 X TRIS BORATE EDTA (TBE)	900mM tris 900mM boric acid 25mM EDTA pH 8.3
GLYCEROL LOADING BUFFER	50% glycerol 10mM EDTA 20mM tris-HCl pH 7.5 bromophenol blue xylene cyanol

2.1.5 E-COLI PROPAGATION

LURIA BROTH	1% Bacto Tryptone 0.5% Bacto yeast extract 0.5% sodium chloride
LURIA-AGAR	1.5% Nobles agar in Luria-broth
LURIA-BERTAINI BROTH	1% Bacto Tryptone 0.5% Bacto yeast extract 1% sodium chloride
LURIA-BERTAINI AGAR	1.5% Nobles agar in Luria-Bertaini broth
TY BROTH	1% Tryptone 1% Bacto yeast extract 0.5% sodium chloride pH 7.4
TY AGAR	1.5% Nobles agar in TY broth
H-TOP AGAR	1% Tryptone 0.8% sodium chloride 0.8% Nobles agar

M9 AGAR	40mM sodium phosphate 20mM potassium chloride 8mM sodium chloride 20mM ammonium chloride 1.5% agarose 20mM magnesium sulphate 0.1mM calcium chloride 0.2% glucose 1mM thiamine-HCl
TRITURATION BUFFER	100mM calcium chloride 70mM magnesium chloride 40mM sodium acetate pH 5.5
SOC	2% Bactotryptone 0.5% yeast extract 10mM sodium chloride 2.5mM potassium chloride 20mM magnesium chloride 20mM magnesium sulphate 20mM glucose

2.1.6 TISSUE CULTURE

MEM	Modified Eagle's Medium
D-MEM	Dulbecco's Modified Eagle Medium 0.001M sodium pyruvate 0.04M sodium bicarbonate
SALINE A	140mM sodium chloride 5mM potassium chloride 5mM glucose
NORMAL SALINE	72mM sodium chloride
TRYPSIN	6% trypsin in saline A
PBS	0.13M sodium chloride 2mM potassium chloride 0.9mM calcium chloride 0.5mM magnesium chloride 11mM sodium phosphate 0.9mM potassium phosphate
PBS'A'	170mM sodium chloride 3mM potassium chloride 10mM sodium phosphate dibasic 2mM potassium phosphate dibasic
PENICILLIN/STREPTOMYCIN	20,000 units per ml of each

FUNGIZONE	0.02% fungizone
BICARBONATE	0.52M sodium bicarbonate 0.001% phenol red
NAPHTHALENE BLACK	0.1% Naphthalene Black 1.36% sodium acetate 6% acetic acid
10 X HBSS	5% HEPES 8% sodium chloride 0.37% potassium chloride 0.125% sodium phosphate
100 X GLUCOSE	10% glucose
10 X DEAE-DEXTRAN	0.5% DEAE-dextran

2.2 TISSUE CULTURE CELLS

HEp-2C — Human Caucasian larynx carcinoma epitheliod cells.
(Toolan, 1954; Moore, *et al.*, 1955)

Growth medium	Plaque assay overlay
1 X MEM	1 X MEM
4% Bicarbonate	5% Bicarbonate
5% FCS	2% FCS
1% P/S	1% P/S
1% Fu	1% Fu
	1% Nobles agar

BGM — African Green Monkey kidney cells.
(Barron, *et al.*, 1970; Dahling, *et al.*, 1974))

Growth medium	Plaque assay overlay
1 X MEM	1 X MEM
2.5% Bicarbonate	5% Bicarbonate
5% FCS	4% FCS
1% P/S	1% P/S
1% Fu	1% Fu
1.5% HEPES	1% Nobles agar

L20B — Mouse connective tissue derived cells transformed with a cDNA clone of human poliovirus receptor. A kind gift from V Racaniello.
(Pipkin *et al.*, 1993)

Growth medium	Plaque assay overlay
1 X D-MEM	1 X MEM
10% FCS	3% FCS
1% Glutamine	5% Bicarbonate
1% P/S	1% P/S
1% Fu	1% Fu
	1% Nobles agar

2.3 *E-COLI* CELLS

TG1 — The TG1 strain was used to transform and grow M13 phage. Its genome included a deletion in the lacZ gene so that blue/white selection could be used with vectors carrying the complement gene. It also included a supE gene required by phage and the F factor that allowed infection by M13 phage. These cells were grown in TY broth.

XL1-BLUE — This E-Coli strain was also used to transform and grow M13 phage in place of the TG1 strain when it lost the ability for M13 infection. The XL1-blue genome included the F factor for M13 infection, a deletion in the lacZ gene for blue/white selection, the recA gene that prevents homologous recombination and the supE gene that is required by phage. These cells were grown in TY broth.

DH5α — The DH5α E-Coli strain was used to transform and grow BR223 plasmid DNA. Its genome included the recA gene that prevent homologous recombination, the hsdR17 gene that abolishes DNA restriction but not protective methylation and also carried the supE gene. These cells were grown in Luria-Bertaini broth or Luria broth when preparing electro-competent cells.

2.4 PRIMERS

13/II — antisense, for types 1,2 and 3 poliovirus, at position 629-645
ACC GGA TGG CCA ATC C

PCR F — sense, for types 1, 2 and 3 poliovirus, at position 23-51
CCA GAG GCC CAC GTG GCG GCT AG

PCR 9 — antisense, for type 3 poliovirus, at position 774-799
GAG CGC CTA CTT TTT GGG ATG ATA C

αLL390 — sense, for type 2 poliovirus, at position 390-404
GCC ACG GGA CGC TAG

LL390 — antisense, for type 3 poliovirus, at position 390-404
CTA GCG TCC CGT GGC

AM20 — sense, for type 2 poliovirus, at position 3204-3225
GAG TGG TGA ATG ACC ACA ACC C

AM21 — antisense, for type 2 poliovirus, at position 4566-4588
AGT GAG TAT GTG GAG GTG TTC TC

AM24 — antisense, for type 2 poliovirus, at position 3559-3582
CTC CTG GAT TCA CAG TAG TAC AC

AM27 — antisense, for type 2 poliovirus, at position 3857-3876
AAT CCA CTC CCA AAT GCA GC

AM31	sense, for type 2 poliovirus, at position 3521-3543 GGC AUA GAC UCA AUU GCU AGA UG
AM37	sense, for type 2 poliovirus, at position 3605-3631 CTT TCA TTG GGC CCN NC/GT TCC AAT ACA
AM38	sense, for type 2 poliovirus, at position 3605-3631 CTT TCA TTG GGC CCA CCN TCC AAT ACA
AM39	antisense, for type 3 poliovirus, at position 459-488 CCT GCT CCA TGG NNN GGA TTA GCC GCA TTC
AM40	antisense, for type 3 poliovirus, at position 459-488 CCT GCT CCA TGG TAG GAT TAG CCG CAT TC
AM41	antisense, for type 3 poliovirus, at position 459-488 CCT GCT CCA TGG AGG ATT AGC CGC ATT C
WTPmlI3266	sense, for coxackie virus, at position 3266-3288 AAA CAC GTG AGA GTG/C TGG TGC/T CC
WTBstE3939	antisense, for coxackie virus, at position 3916-3939 TCT GTA ATG GTG CTG GTN ACC AT
M13 1211	antisense, for M13, at position 6290-6306 GTA AAA CGA CGG CCA GT
All 13	antisense, for types 1, 2 and 3 poliovirus, at position 4153-4163 AGU UGG UUG AA
All 14½	antisense, for types 1, 2 and 3 poliovirus, at position 3809-3821 CAG ACA UUA GAG A
ARWTI	sense, for type 1 poliovirus, at position 3337-3356 GTG/A GAT/C TAT/C AAA/G GAC/T GGN AC
3'WTAR	antisense, for types 1, 2 and 3 poliovirus, at position 3810-3829 ATN CCC/T TGC/T TCC ATN GCC/T TC
ARWTIII	antisense, for type 3 poliovirus, at position 3705-3715 AGT ATG CCA CC
WT#3	antisense, for types 1, 2 and 3 poliovirus, at position 3731-3747 CCA TGT TGA CAC/T CTG AG
WT#4	antisense, for types 1, 2 and 3 poliovirus, at position 3797-3815 CAT TCA AAT CCC TGA TGT C

2.5 METHODS

2.5.1 DNA EXTRACTION

DNA extraction was carried out as described in Sambrook *et al.*, (1989) on three different scales. This was performed on overnight cultures prepared by inoculating medium with a single colony from an agar plate and incubating in an orbital shaker at 37°C. A 'mini' prep used 1ml to produce enough DNA for quick diagnostic purpose where as 'midi' and 'midi/maxi' preps, on 20ml and 200ml of culture respectively, were carried out to produce large amounts of DNA for RNA transfection and cloning.

Initially cells were pelleted and all traces of growth medium removed. The pelleted cells were then resuspended in solution 1, transferred to ice and solutions 2 and 3 were added in ratios of 0.05:0.1:0.075 with respect to the original volume of culture used. With the addition of RNaseA, this was incubated at room temperature for 15 min before precipitated proteins were pelleted and supernatant removed. For a 'mini' prep, phenol/chloroform extraction was carried out before the DNA was precipitated at -20°C in either 1 volume of isopropanol or 2 volumes of ethanol and 0.1 volume of 3M sodium acetate. DNA could then be used in a sequencing reaction. For larger scale preps the DNA was precipitated prior to phenol/chloroform and chloroform extraction. A selective re-precipitation followed this. Concentration of DNA was estimated by visualisation on an ethidium bromide agarose gel.

2.5.2 RESTRICTION DIGESTS

Restriction digests were carried out as suggested by the enzyme manufacturer's protocol.

2.5.3 PCR REACTIONS

A standard 100µl PCR reaction (Saiki *et al.*, 1985) contained 200ng of each primer, 0.2mM dNTP, 1 X manufacturer's buffer, 1 unit of Taq DNA polymerase enzyme (Saiki *et al.*, 1988) and were overlayed with mineral oil to prevent evaporation inside the tube. SuperTaq (Stratech) was used when preparing DNA for sequencing reactions where as AmpliTaq (Perkin Elmer) was selected when preparing DNA for cloning experiments due to its higher fidelity. A Hybaid OmniGene TR3 SM2 thermocyler with an accuracy of ±0.5°C was used. Cycle lengths and temperatures were selected to provide optimum conditions for each set of primers.

2.5.4 VISUALISATION OF DNA

Electrophoresis of DNA was carried on 1% SeaKem agarose gels (Flowgen) in 1 X TBE buffer as electrolyte with bromophenol blue and xylene cyanol as running markers. Gels were freshly made in 1 X TBE with ethidium bromide (Bio-Rad) at 1ng/ml and DNA visualised using UV light (Sambrook *et al.*, 1989).

2.5.5 PURIFICATION OF DNA AND PCR PRODUCTS USING PREP-A-GENE MATRIX

Purification of PCR amplified DNA fragments from PCR reaction mixtures was achieved using Prep-A-Gene silica based matrix (Bio-Rad). The matrix selectively binds DNA fragments of 100 base pairs and above at an efficiency of at least 0.2µg per µl of suspended matrix.

PCR samples were decanted without taking any mineral oil into clean tubes and the DNA allowed to bind to the matrix for 10 min at room temperature in 3.5 volumes of binding buffer. For DNA in agarose slices 6 volumes of binding buffer were used. Matrix was pelleted by centrifuging for 30s, the binding buffer removed by aspiration and washed with

500µl of wash buffer 3 times. The DNA from the dry pellet was eluted in 25µl of GDW at 56°C for 10 min and the matrix separated by centrifugation for 30s.

2.5.6 M13 SINGLE STRANDED PHAGE DNA EXTRACTION

An exponential culture of *E. coli* cells was grown by inoculating TY broth with a 1:100 ratio of an overnight culture and incubating at 37°C for 2 hours in an orbital shaker. Plaques were then picked into 1.2ml of this culture and grown at 37°C for 4.5 hours. Bacterial cells from 1ml of this were then pelleted and discarded. The phage in the supernatant were precipitated at room temperature with 200µl of a 20% PEG 6000 solution in 2.5M sodium chloride. Pelleted phage were resuspended in 100µl of TE buffer and extracted once with phenol mix before precipitating at -20°C in 3 volumes of ethanol and 0.1 volume of 3M sodium acetate. Single stranded DNA was then pelleted, washed in 70% ethanol and dried before eluting in TE buffer.

2.5.7 RNA PURIFICATION FROM TISSUE CULTURE FLUID AFTER PREPARATION OF VIRUS ON SUCROSE GRADIENTS

Continuous 30ml sucrose gradients were prepared in centrifuge tubes from 15ml of the 15% solution and 15ml of the 45% solution. Virus in tissue culture fluid was mixed with 1 drop of Nonidet P40 (BDH) and added carefully to the top of the gradient and centrifuged at 80000XG for 4 hours at 4°C. The gradients were then harvested by bottom puncture and 2ml fractions were collected. From many years of experience the peak fractions 3, 4, 5 and 6 are found to contain poliovirus particles. These were pooled and the virus pelleted by centrifuging at 60000XG overnight. Pelleted virus was resuspended in 500µl SDS/sodium acetate, extracted twice with phenol/ chloroform mix, once with chloroform and precipitated in 2 volumes of ethanol and

50mM sodium chloride at -20°C for 1 hour. The clean RNA was then pelleted, dried and resuspended in GDW.

2.5.8 RNA EXTRACTION FROM TISSUE CULTURE FLUID

Aliquots of 900µl of tissue culture fluid were treated with 1µg Proteinase K (Boehringer Mannheim) in 1% SDS by incubation at 37°C for 15 min. This was followed by extraction with phenol/chloroform twice and chloroform once. RNA was then precipitated in 0.1 volume of 3M sodium acetate and 2 volumes of ethanol at -20°C for 1 hour. Pelleted RNA was washed with 70% ethanol and freeze dried before eluting in GDW. Storage in 70% ethanol at -20°C decreased the rate of RNA degradation. This method was described in Macadam *et al.*, (1989).

2.5.9 RNA TRANSFECTION FROM PLASMID DNA

Infectious RNA was generated from a linearised plasmid which contained a T7 promoter upstream of full length cDNA copies of the poliovirus genome (van der Werf *et al.*, 1986).

Plasmids were linearised with SacI (New England Biolabs) according to manufacturer's protocol and the DNA was cleaned using the Prep-a-Gene method. At least 500ng of DNA was used in a 50µl reaction containing 1 X manufacturer's buffer, 5mM DTT, 0.5mM rNTP's, 15units RNAguard (Pharmacia) and 15units T7 RNA polymerase (Stratech). This was incubated at 37°C for 30 min and a sample run on a 1% agarose gel to check for synthesis of RNA. A mixture of ice cold 1 X HBSS/glucose/DEAE-dextran was added, 200µl for a 25cm^2 flask and 1ml for a plaque assay, and the reaction left on ice for 30 min. Serial log dilutions were made in the 1 X mix for an assay. This was then used to inoculate cells sheets which had been washed 3 times in PBS'A'. Inoculum was removed by aspiration, and cells incubated at room temperature for 20 min. Inoculum was removed by

aspiration, growth medium added and cells incubated at 37°C until CPE was achieved whilst the cell control survived. Cell sheets were then frozen and virus collected as for general virus growth.

2.5.10 cDNA SYNTHESIS

Synthesis of cDNA from tissue culture extracted RNA was carried out using AMV reverse transcriptase (Life Sciences Inc). Reactions were carried out in 0.4mM dNTP, 1 X manufacturer's buffer, 5ng random hexamers (Boehringer Mannheim), 7 units of enzyme and incubated for 1 hour at 42°C.

2.5.11 PRIMER PURIFICATION AFTER SYNTHESIS

Primers were made 'in house' on an Applied Biosystems 392 RNA/DNA synthesiser. The primers were eluted from columns using 1.5ml of a 35% ammonium solution pushed through at 0.5ml every 30 min. Deprotection was carried out by incubating at 56°C overnight and the oligonucleotide precipitated in 0.1 volume 3M sodium acetate and 3 volumes of ethanol at -20°C for 1 hour. The pelleted primer was washed in 70% ethanol, freeze dried and eluted in GDW. Concentration of the primer was measured by optical density (OD) at 260nm on a Philips PU 8720 UV/VIS scanning spectrophotometer using the following equation:

$$OD \times 33 = \text{concentration in } \mu g/ml.$$

2.5.12 T7 DNA SEQUENCING

Chain-terminating sequencing (Sanger *et al.*, 1977) was performed on a number of different DNA starting materials: 1) ds plasmid DNA; 2) ds PCR product and 3) ss M13 DNA with a minimum of 100ng DNA.

1. Ds plasmid DNA was denatured in 0.2M sodium hydroxide at 37°C for 15 min and then precipitated in 0.3 volume of potassium acetate (pH 4.8) and 7.5 volumes of ethanol at

-20°C for 1 hour. DNA was then pelleted and dried before eluting in a total volume of 14µl with 100ng of primer and 2µl of annealing buffer. Annealing took place at 37°C for 20 min.

2. Ds PCR products that had been prepared using the Prep-a-Gene method were denatured with 100ng primer and 0.2M sodium hydroxide at 37°C for 10 min. This was then neutralised with 0.2M hydrochloric acid and 2µl annealing buffer in a total volume of 14µl and incubated at 37°C for 10 min.
3. Ss M13 DNA was simply annealed with 100ng primer and 2µl annealing buffer in a 14µl reaction at 60°C for 10 min.

The denatured/annealed DNA mixes were then labelled with a reaction containing 2µl labelling mix, 1µl P^{32} αdATP (ICN), 1µl 0.1M DTT, 2µl GDW and 3 units T7 DNA polymerase at room temperature for 5 min. This mix was then equally divided into 2.5µl of each termination mix and incubated at 37°C for 5 min before the reaction was stopped by the addition of formamide dye. Samples were boiled before electrophoresis on a 6% acrylamide gel.

2.5.13 CYCLE SEQUENCING

Each sequencing reaction required 2pmol of primer to be labelled with 2pmol (6µCi) P^{32} γdATP using 1unit of polynucleotide kinase (Pharmacia) and 1X manufacturer's buffer. This reaction was incubated at 37°C for 20 min and label incorporation was measured. This was achieved by dotting 0.5µl on adsorbent paper and comparing the counts per s recorded before and after the primer was precipitated with 10% trichloroacetic acid and washed with ethanol. The primer was used in sequencing only when percentage incorporation exceeded 20%. The primer was then heated to 90°C for 2 min to denature the enzyme.

Approximately 200ng of PCR product, cleaned using the Prep-a-Gene method, was needed for each reaction. Initially a 16µl reaction mix was made for each DNA containing 2pmol labelled primer, 1 X manufacturer's buffer and 1 unit of SuperTaq (Stratech). This was then equally divided between the 4 termination mixes to make a total of 8µl and overlayed with mineral oil. Using the Hybaid OmniGene thermocyler 10 cycles were performed with a 40s denaturing step, a 30 s annealing step and a 2 min elongation step. The reaction was then stopped with formamide dye and samples boiled for 2 min before electrophoresis.

2.5.14 RNA SEQUENCING

Sequencing was carried out using a method similar to that described in Evans *et al.*, (1985) using RTase in place of the Klenow fragment of *E Coli* DNA polymerase I. RNA purified by sucrose gradient was incubated at room temperature for 5 min with 15pmol of primer, 1 X manufacturer's buffer, 1µCi P^{32}αdATP, 0.15% labelling mix and 5units RTase (Boerhinger Mannheim). This was then equally divided between 2µl each of the four termination mixes and incubated at 42°C for 10 min. The reaction was stopped by adding formamide dye to each mix and samples were boiled for 2 min before loading on a 6% acrylamide gel for electrophoresis.

2.5.15 POLYACRYLAMIDE GEL ELECTROPHORESIS OF SEQUENCING REACTIONS

Sequencing reactions were run on 6% acrylamide denaturing gels at 15mA per gel, using bromophenol blue and xylene cyanol as approximately 10 and 100 base pair markers respectively. Cambridge gel apparatus was used with eared plates siliconised with dimethyldichlorosilane to repel the gel and non-eared plates treated with 2.5ml wacker solution to make the gel adhere. Plates were taped using x-ray film thick spacers and TBE buffer as running buffer. Once run as required the plates were separated, the gel fixed for

15min in 10% glacial acetic acid, washed for 15 min and dried in an oven for 30 min before exposing to x-ray film.

2.5.16 PLATING M13 BACTERIOPHAGE

Competent TG1 cells were thawed on ice and 100µl used per ligation reaction along with 1.5µl DMSO and 5µl of ligation mix and incubated on ice for 30 min. This mix was then heat shocked at 42°C for 90s before adding to 3ml molten H-top agar containing 200µl of an overnight culture of TG1 cells. The agar was then poured onto TY-agar plates which were inverted and incubated at 37°C overnight. When blue/white selection was used 480µg IPTG and 600µg X-gal was added to the molten agar.

2.5.17 PREPARATION OF ELECTRO-COMPETENT DH5α CELLS

This method was described in the electroporator manual (1991). An overnight culture of cells was prepared by inoculating 500 ml Luria-broth with a single colony from a stock plate, kept at 4°C and re-plated every 2 months. This was then used to inoculate Luria-broth at a ratio of 1:100 and cells were grown at 37°C in an orbital incubator to an optical density of 0.4 at 600nm. After incubating on ice for 1 hour cells were pelleted at 4000XG at 4°C for 15 min and resuspended in 1 volume of ice cold GDW. This was repeated with continually smaller volumes of 0.5, 0.5, 0.02 of GDW and finally a 0.003 volume of filter sterile cold 10% glycerol. Aliquots of 200µl were stored at -70°C.

2.5.18 ELECTROPORATION

For each reaction, 1µl of a 1:5 dilution of ligation mix was added to 40µl of cells and allowed to stand on ice for 1 min. Electroporation was carried out as described in the operating manual (1991) in disposable 2mm gap cuvettes using a BTX ECM 600 Electroporation System at 2.45kV and 129Ω for 5-6 ms. 1ml of SOC was immediately added and cells

incubated at 37°C for 1 hour prior to plating out on selective agar. Conditions were advised in the electroporator manual (1991).

2.5.19 PREPARATION OF CHEMICALLY-COMPETENT TG1 CELLS

An overnight culture of cells was prepared by inoculating TY-broth with a single colony from a stock M9 plate, kept at 4°C and re-plated every 2 months. This was then used to inoculate TY-broth at a ratio of 1:100 and cells were grown at 37°C in an orbital shaker to an optical density of 0.5 at 600nm. After incubating on ice for 2 hours the cells were pelleted at 2500XG at 4°C for 15 min and resuspended in 1 volume (of original culture volume) trituration buffer. After incubation on ice for 45 min this was repeated with resuspension in 0.1 volume. Glycerol was added to a final concentration of 15% and aliquots stored at -70°C.

2.5.20 TEMPERATURE SENSITIVE PLAQUE ASSAYS

Temperature sensitive plaque assays were performed as described in Macadam *et al.*, (1991b). Cell monolayers were prepared in 6 well plates, duplicates were prepared for each virus at each temperature used. After removal of growth medium wells were inoculated with serial log dilutions of virus diluted in 1 X MEM and allowed to adsorb to the cells for 20 min. Cells were then overlayed with a 1% solution of Nobles Agar (Difco) in a 1 X MEM growth medium and sealed with pressure film when set. Inverted plates were then incubated in CO_2 incubators or thermostatically controlled water baths, accurate to within 0.01°C (Grant SP). Temperatures were determined using a Digitron 3204-pt platinum probe digital thermometer. After 3 days, agar was removed and cell sheets stained with Naphthalene Black. The temperature, T°C at which a virus is defined as ts is the temperature at which $\log_{10}$(titre at 35°C/titre at T°C)=1.

2.5.21 VISUALISATION OF PLAQUES

Virus was grown on cell monolayers in 6 well plates as for the ts plaque assays and overlayed with 1% Nobles Agar and 1 X MEM medium. After 3 days incubation a further 1.5ml overlay was added containing 1% Nobles Agar, 0.017% filter sterile neutral red (Sigma) 1 X MEM growth medium and plates were incubated the right way up at 35°C for 4 hours. After incubation, the agar above the visible plaques was taken out using a glass pasteur pipette and incubated in 1ml 1 X growth medium at 4°C overnight so that any virus could diffuse into the medium. Virus bulk could then be grown using this as inoculum. As Neutral Red is very light sensitive plates were shielded as much as possible.

Plaques were selected according to what the experiment involved. If non ts revertants were required, larger plaques at a relatively high temperature were picked. If virus was simply being plaque purified, plaques were picked at 35°C. All plaques selected were individual and spaced out as much as possible. The highest dilution that plaques were visible was usually the one where plaques were selected.

2.5.22 PROPAGATION OF VIRUS

Viruses were grown on HEp-2C monolayers until CPE was established as described in Minor (1980). Cell sheets were then frozen to -20 °C, thawed to release virus and cell debris removed by centrifugation. Long term storage was carried out at -70°C.

2.5.23 MONKEY NEUROVIRULENT TEST

Monkey neurovirulence was tested in a WHO approved procedure (WHO 1983).

CHAPTER THREE

ABILITY OF CODING CHANGES IN THE PROTEASE 2A TO COMPENSATE FOR PROTEIN TRANSLATION INEFFICIENCIES RESULTING FROM 5'NCR MUTATIONS

3.1 INTRODUCTION

At the start of this project ten non ts revertant polioviruses had been characterised that had retained mutations in domain V that destabilised secondary structure but had alternative mutations in the coding region of 2A. These 2A mutations were able to rescue the ts phenotype resulting from translation inefficiencies that were initially due to the domain V mutations. Destabilising domain mutations are usually compensated for *in vivo* by further domain V mutations that re-form the secondary structure necessary for internal initiation of translation of the viral RNA. 2A is a small viral protease responsible for shut off of host cell translation and polyprotein processing but these rescuing mutations appeared to reveal yet another function. The work carried out prior to that described in this thesis was published in Macadam *et al.*, (1994).

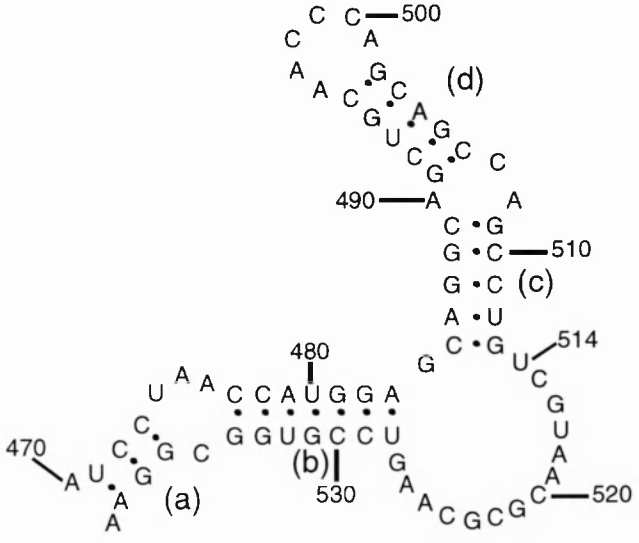


FIGURE 3.1 A representation of the predicted secondary structure of domain V of the 5'NCR. Numbering and sequence used is of a type 3 Leon poliovirus. Stems are labelled for descriptive purposes.

The viruses in this study were derived from two types of ts parental strains. The first was Sabin 2 (S2) which is ts due to the 481 (484 in type 3 numbering, see fig 3.1) G to A mutation, or a virus whose first 491 bases were from a Sabin 2 virus, the rest from a non ts revertant Sabin 2 virus (P117/S5'). The second type were site directed mutants of

Leon/Lansing constructs (LL) where the first 789 bases were derived from a Leon type 3 virus, the rest from a Lansing type 2 virus (Skinner *et al.*, 1989). They were constructed to introduce base pairing or deletions in domain V which are not naturally occurring *e.g.* LL472/537 UC had a UC mismatch and LLΔ483 had a single base deletion at 483 (see fig 3.1) (Skinner *et al.*, 1989; Macadam *et al.*, 1992). The mutants were ts in their growth in BGM cells and this was measured as the ratio of the pfu/ml at 35°C to that at 39°C (see table 3.1).

PARENT (PRECURSOR)	Log_{10} (pfu at 35°C/ pfu at 39°C)	REVERTANT VIRUS	Log_{10} (pfu at 35°C/ pfu at 39°C)
(P2/117)	0.2	-	-
(P2/1175')	0.3	-	-
P2/Sabin	2.2	S2/39A	0.6
P117/S5'	2.3	S5'39A	0.4
		S5'39B	0.1
		S5'39C	0.1
		S5'39D	0.2
Leon/Lansing (LL)	0.1	-	-
LL472/537UG	2.0	UG/39A	0.3
LLΔ472	4.0	Δ472/39A	0.5
LL479/532UC	2.8	UC/39A	0.2
LLΔ483	4.0	Δ483/39A	0.9
LL514A	3.0	514/39A	0.4

TABLE 3.1 Comparison of ts phenotypes of parent and revertant viruses in BGM cells. Taken from Macadam *et al.*, (1994).

All parent viruses were ts to different degrees in BGM cells (see table 3.1) as the changes to domain V disrupt to different extents. For comparison, the equivalent non ts viruses were constructs of Sabin 2 with a 481 G (P2/117 and P2/1175') (Macadam *et al.*, 1991b) and the original Leon/Lansing virus containing no 5'NCR mutations. Revertant viruses were picked following plaque formation at 39°C on BGM cells and were found to be

relatively non ts but again to slightly different degrees from each other (table 3.1). On sequencing, the 5'NCRs of these revertants were identical to the mutant viruses from which they were derived. Further sequencing revealed that all revertants had mutations in the gene for the protease 2A resulting in coding changes (see table 3.2).

ISOLATE	AMINO ACID DIFFERENCE
S2/39A	96 His-Tyr
S5'39A	19 Tyr-His
S5'39B	8 Ala-Val
S5'/39C	79 Thr-Ala
S5'/39D	122 Ile-Val
UG/39A	25 Glu-Gly
Δ472/39A	80 Phe-Leu
514/39A	10 Tyr-Cys
UC/39A	82 Tyr-His
Δ483/39A	23 Thr-Ile

TABLE 3.2 Amino acid differences in the 2A coding regions of non ts revertants at the start of the project. Taken from Macadam *et al.*, (1994).

Protein labelling experiments in BGM cells confirmed that the non ts revertants carrying 2A changes had restored protein synthesis efficiencies comparable to the restoration obtained by direct back mutation in the domain V. In addition there was no apparent difference in shut off of host cellular translation in cells infected with these and other viruses and no differences in levels of extracted viral RNA.

Reconstruction of three of these revertants: 8 Ala-Val; 19 Tyr-His and 96 His-Tyr, by insertion of a PCR amplified 2A coding region into a Sabin 2 clone confirmed the effect of these mutations. All three were able to rescue the ts phenotype brought about by the 5'NCR disruption leading to the conclusion that the changes in 2A were responsible for rescuing this phenotype.

The objective of the work described in this chapter was to identify further 2A changes in non ts isolates so as to detect any clustering, link specific mutations to specific 5'NCR disruptions, to determine any importance of RNA sequence and to gain some idea of the functional basis of the suppressor effects.

RESULTS

3.2 DETECTION OF MUTATIONS IN 2A

A number of non ts revertant viruses that had been plaque picked in BGM cells grown at 39°C or 39.3°C were available for sequencing throughout their 2A genes and virus stocks were grown in 25cm^2 flasks of HEp-2C monolayers at 35°C. This exerts little selection pressure on the virus because in these cells the ts phenotype is observed at much higher temperatures (Macadam *et al.*, 1992), making them more permissive to the phenomenon. Sequencing was achieved by phenol/chloroform extraction of RNA from tissue culture fluid, cDNA synthesis and PCR amplification of the 2A region using primers AM20 and AM21. PCR products were purified using the Prep-a-Gene silica matrix method and the entire gene was sequenced by cycle sequencing using γP^{32}-dATP labelled primers AM27 and AM20. Examples of results from sequencing gels can be seen in fig 3.2. Most had already been sequenced through domain V of the 5'NCR to verify they retained parental disruptions but where this information was not available the area was amplified from cDNA by PCR using primers PCR F and 13/II (see chapter two). Sequencing was carried out as for the 2A region using labelled 13/II as primer.

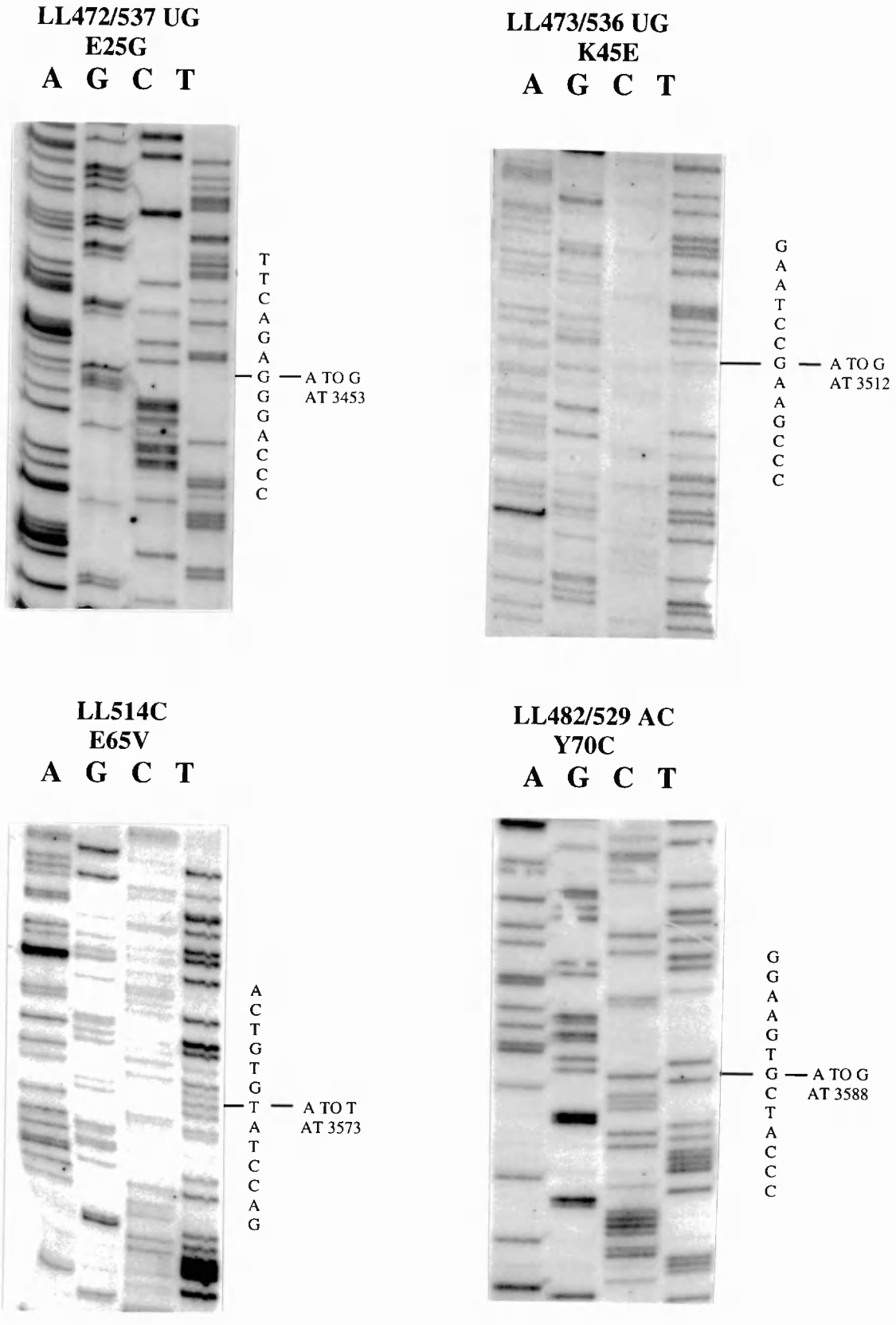


FIGURE 3.2 Photographs of examples of sequencing autoradiographs showing some of the coding changes in 2A. Changes were found in non ts revertant viruses selected in BGM cells. Viruses still contained a mutation in domain V of the 5'NCR as indicated.

3.3 POSITIONS OF 2A CHANGES

Eighteen more BGM selected non ts viruses were sequenced throughout the 2A gene and these data along with previous results published in Macadam *et al.*, (1994) are presented in table 3.3. The titres at the higher temperature of 39°C were chosen as representative so as to display a clear difference between non ts Leon/Lansing and most other ts viruses. It shows that a very broad range of 2A changes were found and no apparent link can be immediately seen. For example, a single 5'NCR disruption was found to give rise to many different mutations in 2A *e.*g. five separate 2A changes were found in Sabin 2 revertants and five different 2A changes were found in revertants of LL472/537 UG. All grew to slightly different degrees at the higher temperature, indicating that the mutations do not affect the protein to exactly the same degree. All revertants displayed similar plaque morphology, generally larger than parental ts viruses but comparable to equivalent non ts viruses. In addition table 3.3 does not convey that some non ts revertant viruses containing the same coding change in 2A were plaque picked on two separate occasions.

5'NCR DISRUPTION	2A CODING CHANGE	Log_{10}(pfu at 35°C/pfu at 39°C)
Sabin 2 - 484 A or P117/S5'	None 8 Ala - Val 19 Tyr - His 79 Thr - Ala 96 His - Tyr 122 Ile - Val	2.2 0.1 0.4 0.1 0.6 0.2
LL471/538 AA	None 65 Glu - Lys	1.7 0.5
LL471/538 GA	None 122 Ile - Val	1.8 0.7
LL472/537 UG	None 25 Glu - Gly 33 Ile - Val 93 Tyr - His 106 Pro - Ser 134 Ser - Thr	2.0 0.3 0.2 0.9 0.5 0.5
LL472/537 UA	None 30 Ala - Pro	1.0 0
LLΔ472	None 80 Phe - Leu 17 Cys - Tyr	4.0 0.5 0.7
LL473/ 536 UG	None 45 Lys - Glu 79 Thr - Ala	2.0† 0.5† 0.4†
LL479/532 UC	None 82 Tyr - His	2.8 0.2
LL482/529 AC	None 25 Glu - Gly 70 Tyr - Cys	2.8 0.3 0
LLΔ483	None 23 Thr - Ile	4.0 0.9
LLΔ483/528	None 48 Gly- Asp	2.5 0.2
LL514A	None 10 Tyr - Cys 19 Tyr- Cys	4.0 0.4 1.5
LL514C	None 65 Glu - Val	1.5 0.2

TABLE 3.3 Final list of non ts revertants of viruses with eleven different 5'NCR disruptions with corresponding amino acid differences in the 2A coding region and comparison of titre at 35°C to that at 39°C in BGM cells.
† Figures are quoted from an assay where the higher temperature was 39.3°C to better show the phenotype of the parent virus.

The changes were found along the entire length of the 149 amino acid 2A polypeptide and substitutions were found to be acidic, basic or hydrophobic with no conservation of size or electrodensity (see table 3.4). A single 2A mutation was found in all viruses sequenced, suggesting that all revertants were using the same method to rescue translation inefficiencies. Changes did not derive from PCR errors as different PCR reactions from the same virus were found to have the same change. In addition a single 2A amino acid change was found in two different viruses *e*.g. a change of threonine to alanine at residue 79 of the 2A polypeptide was found in non ts revertants of Sabin 2 and LL473/536 UG. In all of these cases both viruses were found to have the same base change in the codon despite being selected from different parental viruses None of the RNA sequence changes in 2A were found to be silent and two different substitutions were found at residues 19 and 65 albeit in different parental viruses.

Interestingly five of the mutations changed a tyrosine residue to either a histidine or cysteine residue. Similarly four of the changes resulted in valine residues. The significance of this is unclear but in most of these cases the base changes involved a transition rather than a transversion. Evidence exists to suggest that transitions, changes of A to G or G to A, are more frequently found in polioviruses than transversion, changes of C to U or U to C (Kuge *et al.*, 1989).

2A CODING CHANGE	5'NCR DISRUPTION
8 Ala - Val	Sabin 2
10 Tyr - Cys	LL514A
17 Cys - Tyr	LLΔ472
19 Tyr - His	Sabin 2
19 Tyr - Cys	LL514A
23 Thr - Ile	LLΔ483
25 Glu - Gly	LL472/537 UG LL482/529 AC
30 Ala - Pro	LL472/537 UA
33 Ile - Val	LL472/537 UG
45 Lys - Glu	LL473/536 UG
48 Gly - Asp	LLΔ453/528
65 Glu - Lys	LL471/538 AA
65 Glu - Val	LL514C
70 Tyr - Cys	LL482/529 AC
79 Thr - Ala	Sabin 2 LL473/536 UG
80 Phe - Leu	LLΔ472
82 Tyr - His	LL479/532 UC
93 Tyr - His	LL472/537 UG
96 His - Tyr	Sabin 2
106 Pro - Ser	LL472/537 UG
122 Ile - Val	Sabin 2 LL471/538 GA
134 Ser - Thr	LL472/537 UG

TABLE 3.4 List of differences found in the 2A coding regions of non ts revertants with corresponding 5'NCR disruptions in domain V.

3.4 POSITIONS OF AMINO ACID CHANGES ON PROTEIN MODEL

The alignment depicted in figure 3.3 below shows a comparison of six different poliovirus amino acid sequences obtained from the GenBank database and is reproduced from Macadam *et al.*, (1994). From this, the amino acid sequence of poliovirus 2A was estimated to be approximately 93% conserved. In figure 3.3 the residues where

compensating mutations were found are shown in red. Only four out of twenty are found in variable positions (25, 33, 45 and 134), the rest are at highly conserved positions within polioviruses. It is interesting to note that the change at residue 33 of valine to isoleucine found in a virus with a UG mismatch at 472/537 is also found in Lansing, Leon and Sabin 3 and that the change found at residue 134 in LL472/537 UG is also found in Sabin 1. This would suggest that these 2A changes have no effect on neurovirulence. However changes at 25 and 70 were associated with, but not proved to be, determinants of mouse neurovirulence in a mouse adapted strain LS-a (Lu *et al.*, 1994). Figure 3.3 also shows that the majority of changes are found in the first 100 residues of the protein, with most of these appearing in the first 50. In addition the clustering of changes between residues 8 - 25 and 79 - 96 originally described in Macadam *et al.*, (1994) is not completely consolidated with the additional changes reported here.

```
                1                                                    50
P1/Sabin        ---------- ---------- ---------- -n-------- -t--r-----
P1/Mahoney      ---------- ---------- ----d----- -n---s---- -t--r-----
P2/Sabin        ---------- ---------- ---------- ---------- -a--r-l---
P2/Lansing      ---------- ---------- ---------- -ni--i---- -v------i-
P3/Leon         ---------- ---------- ---k-----t --i------- -v--------
P3/Sabin        ---------- ---------- ---k------ --i------- -v--------
P3/Finland      ---------- ---------- ---------- ---------- -t------i-
Consensus       GFGHQNKAVY TAGYKICNYH LATQEDLQNA VSVMWNRDLL V-ESKAQGTD

                51                                                   100
P1/Sabin        ---------- ---------- ---------- ------n--- ----------
P1/Mahoney      ---------- ---------- ---------- ------n--- ----------
P2/Sabin        -----s--t- ---------- -----i---- ------e--- ----------
P2/Lansing      -------ht- ---------- -----t---- ------e--- ----------
P3/Leon         ---------- ---------- ---------- ---------- ----------
P3/Sabin        ---------- ---------- ---------- ---------- ----------
P3/Finland      -------st- -------sr- ---------- ---------- ----------
Consensus       SIARCNCNAG VYYCESRRKY YPVSFVGPTF QYMEANDYYP ARYQSHMLIG

                101                                                  149
P1/Sabin        ---------- -----h---- ---------- ---t------ ---------
P1/Mahoney      ---------- -----h---- ---------- ---------- ---------
P2/Sabin        ---------- ---------- ---------- ---------- ---------
P2/Lansing      ---------- ---------- ---------- ---------- ---------
P3/Leon         ---------- ---------- --v------- ---------- ---------
P3/Sabin        ---------- ---------- --v------- ---------- ---------
P3/Finland      ---------- ---------- ---------- ---------- ---------
Consensus       HGFASPGDCG GILRCQHGVI GIITAGGEGL VAFSDIRDLY AYEEEAMEQ
```

FIGURE 3.3 Amino acid sequences of poliovirus 2A. Sequences were from the GenBank database and amino acid sequence obtained using the 'translate', 'pileup' and 'pretty' programs on GCG. Conserved residues are shown as dashes (-) and differences to the consensus are shown as lower case letters. Positions where compensating amino acid changes were found are shown in red. Adapted from Macadam *et al.*, (1994).

As discussed in the introduction a putative α-carbon chain model of the protease 2A was constructed based on the structure of bacterial serine proteases (Bazan & Fletterick 1988; Yu & Lloyd 1991). This model does not necessarily represent the definitive folding of 2A but highlighting the positions of the compensating mutations found here in this model points to some evidence of clustering at the top and bottom-left (see fig 3.4) although the effect is not exclusive.

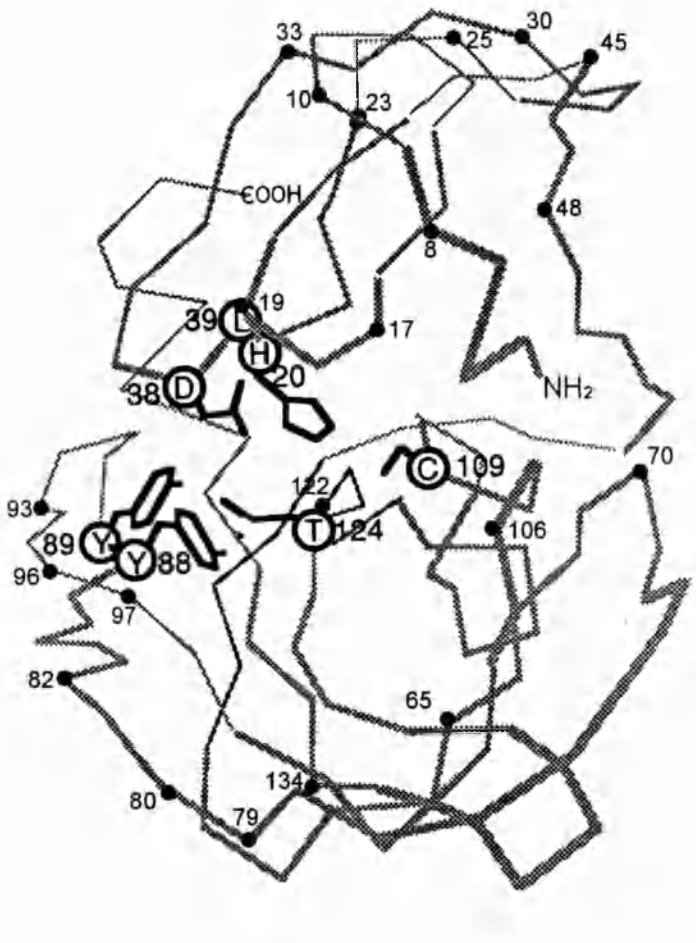


FIGURE 3.4 Possible α-carbon chain model of 2A. The putative catalytic triad is made up of residues 20,38 and 109 and highly conserved residues are at 39, 88, 89 and 124. Probable positions of suppressor mutations are indicated as black circles with residue number.

All the mutations seem to lie on the outside of the protein which could implicate a surface or binding function but this may be simply because 2A is very small. None were found at positions of high conservation in small serine proteases but three mutations, at positions 17,19 and 122 seem to be close to the proposed active site. A proper evaluation of the positions and implications to folding of these changes will require a crystal structure of 2A.

3.5 DIFFERENT EFFECTS OF 2A MUTATIONS WITH A SINGLE DOMAIN V MUTATION

From table 3.4 many different changes in 2A appear to be able to rescue the ts phenotype resulting from a domain V mutation. The amino acid changes show no consistency and can

be found in different positions in 2A. For example a non ts revertant of LL472/537 UG selected in one plaque was found to have sequence heterogeneity and this virus was further plaque purified in HEp-2C cells to separate any sub-populations. A total of four individual plaques were picked, grown and sequenced through domain V to ensure they still had the mismatch and through 2A as described above. Each was found to have a different 2A coding change and on assaying in BGM cells all were less ts than the parent virus.

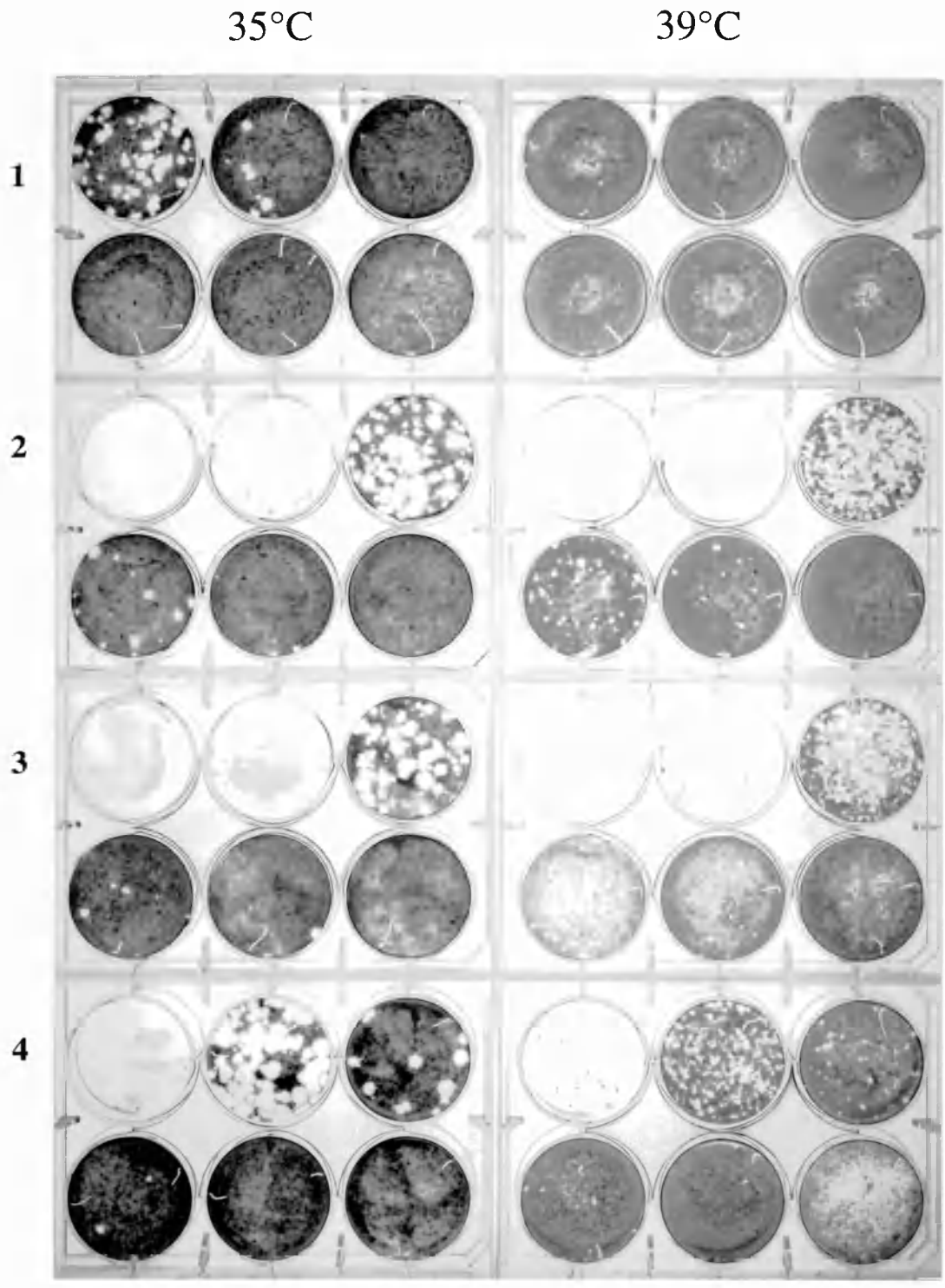


FIGURE 3.5 Plates from a ts plaque assay in BGM cells. The viruses assayed were all Leon/Lansing constructs and had a UG mismatch at 472/537. Three viruses also had amino acid substitutions in 2A: 1 = LL472/537 UG; 2 = LL472/537 UG E25G; 3 = LL472/537 UG S134T; 4 = LL472/537 UG P106S. Plates incubated at 35°C in the first column were inoculated with 10 fold dilutions from 10^{-3} to 10^{-8}. Plates incubated at 39°C in the second column were inoculated with dilution from 10^{-2} to 10^{-7}.

Figure 3.5 depicts some of the plates from the ts assay in BGM cells for these viruses. It clearly shows the dramatic drop in virus growth of the parent virus at the higher temperature of 39°C and the effect of the 2A coding changes on virus growth. Comparison of plaque morphology shows a reduction in size at the higher temperature which is a general trend in all cases but that all viruses were comparable at 35°C. Although the revertants all grew at the elevated temperature they were non ts to slightly different degrees as can be seen from fig 3.5 and table 3.5 further indicating that not all the 2A changes are exactly equivalent.

Virus	2A Coding Change	Log_{10} (pfu at 35°C/pfu at T°C) for BGM cells			
		38°C	38.5°C	39°C	39.5°C
Leon/Lansing	-	0	0.2	0.4	0.75
472/537 UG	-	0.5	1.7	2.7	3
472/537 UG/2A.1	25 Glu - Gly	0	0.2	0.4	1
472/537 UG/2A.2	134 Ser - Thr	0	0.3	0.5	0.6
472/537 UG/2A.3	93 Tyr - His	0	0.2	0.4	1
472/537 UG/2A.4	106 Pro - Ser	0	0.3	0.9	0.6

TABLE 3.5 Comparison of ts phenotype in BGM cells of a non ts revertant and subsequent plaque purified viruses containing 2A coding changes with precursor viruses.

The graph in fig 3.6 also shows that all the non ts revertants of LL472/537 UG with 2A mutations have similar profiles to the Leon/Lansing (472/537 CG) virus and reductions in titres are very small with increasing temperature. In contrast, the LL472/537 UG virus has a large reduction in titre with temperature.

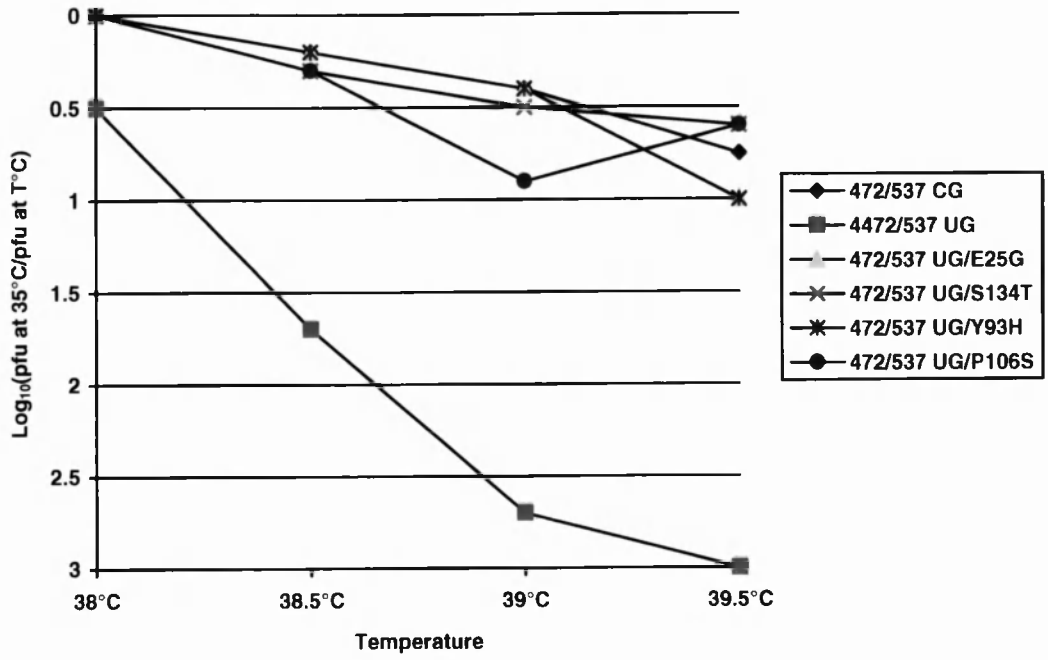


FIGURE 3.6 Graph of log reduction in titre with increasing temperature for Leon/Lansing, LL472/537 UG and non ts revertants of it carrying 2A mutations.

3.6 REVERSION OF TS IN BGM CELLS

There were a very large number of non ts revertant viruses that had previously been selected in BGM cells and found to retain 5'NCR mutations and consequently not all of them were sequenced during this project to confirm they had a change in 2A. In addition, for many domain V mutations, *e.g.* deletion mutants, there is no direct back mutation possible. It is very difficult therefore to calculate the proportion of revertants that are likely to have mutations in 2A rather than direct back mutations in domain V. However there are seven viruses in table 3.3 able to potentially revert in the 5'NCR but this reversion was never observed in BGM cells. Seventeen non ts revertants of these viruses were found to have 2A changes. For these viruses it would appear that the probability of reversion via changes in 2A was at least seventeen times greater than the probability of reversion in domain V.

The fact that there were so many revertant viruses that had 2A mutations could be an indication that there are more places in 2A that the virus can use for reversion. Clearly if

the virus were limited only to reversion in domain V there are two positions where spontaneous mutations can arise and result in a viable non ts virus *i.e.* direct back mutation or a mutation in the opposite strand. If the virus also has the whole of 2A for this, there are more positions where a change could arise during replication. From the data presented above there are at least 22 residues, possibly more, that can potentially mutate. In addition different changes at each position could also occur.

Non ts revertant viruses with domain V reversions can be selected in BGM cells. In particular some viruses with 514 C or A were found to revert back to 514 U (personal communication, A. J. Macadam) and viruses with mismatches in the base pair 472/537 were found to revert only by re-forming the pair (Macadam *et al.*, 1992). However mismatches at 472/537 resulted in the most severe ts phenotype of all viable mutants constructed (Macadam *et al.*, 1992). It is conceivable that the mechanism by which 2A changes are able to compensate for domain V mutations has limits and therefore no 2A changes would ever be found in viruses of this type. This idea will be addressed below.

3.7 EFFECT OF A CHANGE OF PHENYLALANINE TO LEUCINE AT POSITION 80 IN A LL472/537AC VIRUS

The virus LL472/537 AC was made by site directed mutagenesis of a Leon/Lansing clone (Macadam *et al.*, 1992) and has a mismatch of AC at 472/537. This mismatch was found to affect the growth of the virus quite significantly leading to a ts phenotype at the lowest temperature of 36.5°C along with LL472/537 AG and LL472/537 CC, of the viruses discussed in Macadam *et al.*, (1992). The domain V structure of this virus did not appear to support efficient translation.

There were no viruses with compensatory 2A changes from any of the very ts viruses mentioned above. All non ts revertants obtained were found to have re-formed base pairing in domain V relatively quickly as the growth of the virus was so impaired. It was therefore of interest to determine if a change in 2A was capable of rescuing such a phenotype. Using a clone of LL472/537 AC, an amino acid change at residue 80 in 2A was introduced. A change of phenylalanine to leucine at this position was found in a non ts revertant of a virus with a deletion at 472, LLΔ472 (see table 3.3). A deletion at 472 results in a highly ts virus although not as ts as LL472/537 AC. The temperature of expression of the ts phenotype was reported as 37.5°C for LLΔ472 and 36.5°C for LL472/537 AC in Macadam *et al.*, (1992). This could be due to the fact that a deletion at 472 still allows a two base pair stem to form whereas a mismatch here would completely disrupt it. The reconstruction of such a ts virus with a mutation at residue 80 would investigate how well a change in 2A was able to suppress the effects of these 5'NCR changes. In a non ts revertant of LLΔ472 an alternative change of cysteine to tyrosine at residue 17 of 2A was also found.

3.7.1 CONSTRUCTION OF THE LL472/537 AC/F80L VIRUS

A portion of the 2A gene from the virus with the change at residue 80 was amplified in a PCR reaction using primers AM20 and AM27 (see chapter two). Amplified DNA of 672 bases spanned the codon for residue 80 as well as unique restriction sites XhoI at 3297 and BglII at 3790. PCR was carried out using a large amount of DNA, a highly processive and accurate polymerase, AmpliTaq (Perkin Elmer), and as few cycles as possible so as to minimise the likelihood of mistakes being randomly introduced. The DNA was purified using the Prep-a-Gene method prior to cleavage with XhoI (New England Biolabs) and BglII (New England Biolabs) to give a 493 base pair fragment. Plasmid DNA purified

from a liquid culture of the clone containing a cDNA copy of the Leon/Lansing virus with an AC mismatch at 472/537 was also cleaved with these enzymes.

The 493 base pair fragment of 2A with the change at residue 80 and the large fragment of the LL472/537 AC clone were purified from agarose gel slices following electrophoresis of digested DNA by the Prep-a-Gene method. After ligation of the corresponding fragments, DNA was electroporated into DH5α electrocompetent cells. Colonies growing on selective agar were picked and grown in liquid culture. DNA extracted from these was used in a T7 sequencing reaction with primers AM20 and AM27. The entire insert was sequenced to check both the sequence at the codon for residue 80 and to check no other mutations had been introduced by PCR error.

Plasmid DNA from correct clones was purified from liquid culture and linearised using SalI (New England Biolabs). RNA made using T7 RNA polymerase and diluted in a 1X HBSS/glucose/DEAE-dextran mix was used to transfect monolayers of HEp-2C cells in $25cm^2$ flasks. Flasks were incubated at 35°C and checked daily for CPE. RNA is very unstable and prone to degradation. Transfection of poor quality RNA will not result in growth of any virus and can lead to viruses being falsely deemed non viable. In order to discount this, the transfection of the clone was carried out in duplicate with positive and negative controls included. As expected, all cells displayed CPE after only 36 hours except those mock infected. This meant that the LL472/537 AC/F80L virus was viable. Negative control cells were still healthy after six days.

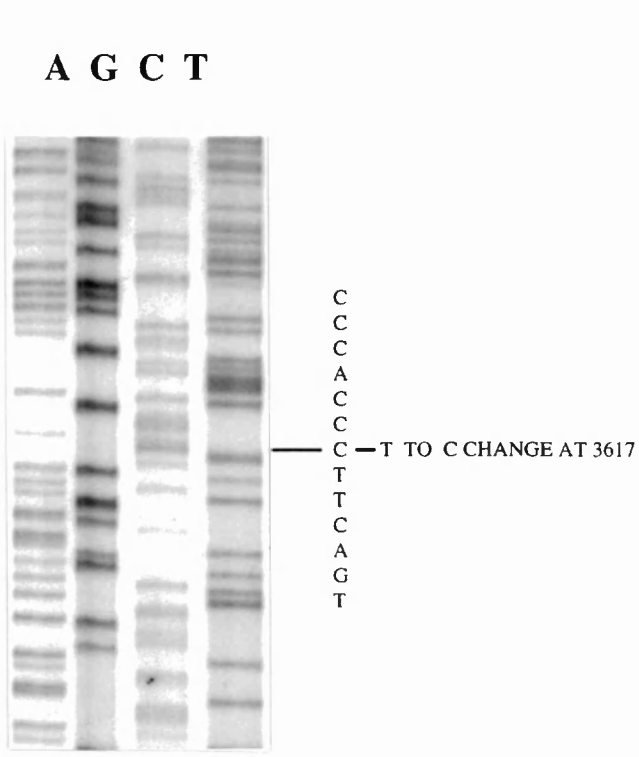


FIGURE 3.7 Photograph of a sequencing autoradiograph of the LL472/537 AC/F80L clone showing the T to C change at 3617. This changes the codon for residue 80 from CUU (phenylalanine) to CCU (leucine). T7 sequencing was carried out on a PCR product made from primers AM20 and AM 27 using AM27 as primer. Lanes are marked AGCT.

Supernatant from flasks of transfected cells, frozen when CPE was observed, formed viral stocks. Once again, sequence of 2A was checked. This was performed on a PCR product using cDNA made from extracted RNA. PCR primers were AM20 and AM27 and AM27 was used in a T7 sequencing reaction (see fig 3.7). The virus was also assayed in BGM cells with Leon/Lansing and LL472/537 AC as controls. Results show that the change of phenylalanine to leucine at residue 80 had no effect on the ts of the virus (see table 3.6). Although the domain V sequence was not obtained there was no evidence or either a change to GC or AU at this base pair as both would be reflected in a smaller drop in titre with temperature. Viruses with both the 472/537 AC mismatch and the change at residue 80 of 2A were as ts as the parent virus LL472/537 AC. It must be noted that if the values for titre reduction at 39°C in table 3.3 are compared the virus LL472/537 AC does not

appear to be very ts. The values at 38°C display ts a little more. This highlights the inter-assay variation that can occur and the need for control viruses in each assay.

VIRUS	Log_{10}(pfu at 35°C/pfu at 38°C)	Log_{10}(pfu at 35°C/pfu at 39°C)
Leon/Lansing	0.4	1.7
LL472/537 AC	1.8	2.6
LL472/537 AC/F80L	2.1	2.5

TABLE 3.6 Table to show the log drop in titre at 38°C and 39°C with comparison to that at 35°C of the viruses Leon/Lansing, LL472/537 AC and LL472/537 AC/F80L in BGM cells. Results shown for LL472/537 AC/F80L are an average of the two viruses reconstructed separately from the same original clone.

These results suggest that not all ts phenotypes resulting from changes in domain V of the 5'NCR can be compensated for by coding changes in the protease 2A. A change of phenylalanine to leucine at residue 80 was able to compensate for the strong ts determinant of a deletion at 472 but was unable to compensate for an even stronger ts determinant of an AC mismatch at 472/537. Although only one change was reconstructed here it would appear that there is a limit to the compensatory role of 2A with a single amino acid change. Coupled with this is the fact that no non ts revertants of this virus were found with changes in 2A. Obviously to test this, other 2A mutations should be introduced into this clone as not all of the substitutions associated with a single domain V mutation are equivalent. It is possible therefore that other changes may have a stronger effect.

3.8 INTRODUCTION OF CHANGES IN AMINO ACID AT POSITIONS 79 AND 80 OF THE PROTEIN 2A

Further evidence that all the changes in 2A listed in table 3.3 were responsible for rescuing the ts phenotype can be provided by cloning reconstruction. However three different 2A mutations had already been shown by this method to suppress a ts phenotype resulting from

a domain V mutation in previous work (Macadam *et al.*, 1994) and all non ts revertants studied here proved to bear a change in 2A as described above. Furthermore no silent changes were discovered so it would appear that the actual change in amino acid was more important than the change in RNA. Despite these arguments it is conceivable that the changes in RNA of the 2A gene either enhanced or inhibited secondary or tertiary structure which could be responsible for the enhancement of translation efficiency.

This question was addressed using site directed mutagenesis to introduce a range of mutations in a single residue of 2A. Codons for residues at positions 79 and 80 of the protease 2A were altered so as to bring about different amino acids and non synonymous codons for the same amino acid at these positions. A change of threonine to alanine at position 79 was found in non ts revertants of both Sabin 2 and LL473/536 UG and a change of phenylalanine to leucine at position 80 had been found in a non ts revertant of LLΔ472 (see table 3.3). Reconstruction described here was carried out using a cDNA clone of Sabin 2. For ease of manipulation, the ApaI restriction enzyme site at position 3522 was previously removed by cloning a silent change into the RNA at this point that is not detrimental to the virus (personal communication, A. J. Macadam). Absence of this site naturally occurs in Sabin 1 and its removal leaves a single ApaI site that can be used for cloning purposes. In this experiment unique restriction sites for ApaI and SnaBI restriction enzymes were utilised at positions 3617 and 4455 respectively.

Degenerate primers AM37 and AM38 (see chapter two)were designed for this mutagenesis both spanning the ApaI site. For AM37, the bases corresponding to the 79th codon of 2A were degenerate *i*.e. NNC/G. This meant that a combination of bases could be introduced into the amplified DNA so that any amino acid could result from the full length clone. In the case of AM38, only the first base of the 80th codon was degenerate. For all three to be

degenerate the primer would have been longer to incorporate the ApaI site and the necessary overlap for binding to the template. Consequently only isoleucine, valine, leucine or phenylalanine could be introduced at position 80 from codons ATC, GTC, CTC or TTC.

Part of the 2A gene was amplified so as to incorporate the ApaI and SnaBI sites using 2µg of DNA of the clone with either primers AM37 and AM21 or AM38 and AM21. Ten cycles of a PCR reaction generated DNA containing a mixture of sequence at the codons for amino acids 79 (from AM37) or 80 (from AM38). Amplification was carried out using 1u of AmpliTaq (Perkin Elmer), a highly processive polymerase and mixed with 0.1u Vent DNA polymerase (New England Biolabs), a proof reading enzyme, so as to obtain a high quantity of DNA with as few mistakes as possible (Barnes 1994). This mixed DNA was ligated into the original Sabin 2 clone using restriction enzymes ApaI and SnaBI.

After transfection into E Coli cells separate colonies were selected and grown in liquid culture. DNA from these was then sequenced using primer All 14½ in a T7 sequencing reaction. From this, for position 79, ten clones were selected with ten different codons coding for eight different amino acids, six of which were not previously found at this position. This meant that there were two clones with alanine and two clones with serine arising from different codons. For position 80, four clones were selected with four different codons for amino acids although one of them with a phenylalanine at position 80 was identical to Sabin 2. This is summarised in table 3.7. After transfection with RNA transcribed from these clones, resulting viruses were assayed in BGM cells to determine whether the changes introduced acted in the same way as the changes previously identified in the non ts revertants.

CLONE	CODON SEQUENCE	AMINO ACID
SABIN 2	ACC/TTC	79-Thr/80-Phe
S2/2A/79E	GAG/TTC	79-Glu/80Phe
S2/2A/79R	AGG/TTC	79-Arg/80Phe
S2/2A/79W	TGG/TTC	79-Trp/80Phe
S2/2A/79L	CTG/TTC	79-Leu/80Phe
S2/2A/79Ta	ACG/TTC	79-Thr/80Phe
S2/2A/79H	CAC/TTC	79-His/80Phe
S2/2A/79Aa	GCG/TTC	79-Ala/80Phe
S2/2A/79Ab	GCC/TTC	79-Ala/80Phe
S2/2A/79Sa	TCG/TTC	79-Ser/80Phe
S2/2A/79Sb	AGC/TTC	79-Ser/80Phe
S2/2A/80F	ACC/TTC	79-Thr/80-Phe
S2/2A/80V	ACC/GTC	79-Thr/80-Val
S2/2A/80L	ACC/CTC	79-Thr/80-Leu
S2/2A/80I	ACC/ATC	79-Thr/80-Ile

TABLE 3.7 List of mutant viruses made by site directed mutagenesis of a Sabin 2 clone. Changes introduced into the RNA sequence are listed along with the resulting amino acid change at either position 79 or 80 in 2A.

3.8.1 EFFECTS OF AMINO ACID AT RESIDUE 79 OF 2A

Ten viruses with differences in the 79th codon of 2A were assayed in BGM cells along with the original non ts revertant of Sabin 2 with a threonine to alanine substitution at this residue. The results are displayed in table 3.8. From these, the original non ts revertant of Sabin 2 (Sabin 2/T79A) is shown to be more ts than the Sabin 2 virus with a G at 481. This further illustrates that a change in 2A is not necessarily equivalent to re-forming domain V secondary structure in terms of effect on a ts phenotype.

VIRUS	Log_{10}(pfu at 35°C/pfu at 37.5°C)	Log_{10}(pfu at 35°C/pfu at 38.5°C)
Sabin 2 - 481 G	0.3	0.6
Sabin 2/T79A	0.3	1.0
Sabin 2	0.1	>2
S2/2A/79E	0.1	1.5
S2/2A/79R	0.8	2.9
S2/2A/79W	0	0.8
S2/2A/79L	0.2	2.7
S2/2A/79Ta	1.2	3.2
S2/2A/79H	0	0.7
S2/2A/79Aa gcg	0.2	1.4
S2/2A/79Ab gcc	1.1	1.8
S2/2A/79Sa tcg	0.3	1.3
S2/2A/79Sb agc	0.1	2.7

TABLE 3.8 Table of results from a ts assay in BGM cells of Sabin 2 viruses with different amino acids at residue 79. The log drop in titre was calculated at 37.5°C and 38.5°C from that at 35°C. A non ts Sabin 2 virus with a G at 481 was used as comparison along with the original revertant with a change at residue 79.

It is also clear that not every amino acid substitution at this position has the same effect. Some of the ts viruses, S2/2A/79E, S2/2A/79L, S2/2A/79Aa and S2/2A/79Sb, are more like the non ts viruses at lower temperatures but very ts at the higher temperature. This would imply that these amino acids are able to exert only a limited effect. Other viruses, S2/2A/79R, S2/2A/79Ta and S2/2A/79Ab already display a relatively ts phenotype at 37.5°C. The non ts viruses are S2/2A/79W, S2/2A/79H and to some extent S2/2A/79Sa. None of the substituted amino acids in these three groups of viruses show any similarity in size, shape or electrodensity *e.g.* basic arginine, polar threonine and non polar alanine at residue 79 result in ts viruses.

Viruses with different codons for the same amino acid were also represented in this experiment. These synonymous amino acids were alanine, threonine and serine with an alanine present in Sabin 2 and threonine present in the original non ts revertant selected in BGM cells. Table 3.9 below compares these viruses and indicates the drop in titre with temperature detected in BGM cells.

VIRUS AND CODON	Log_{10}(pfu at 35°C/pfu at 37.5°C)	Log_{10}(pfu at 35°C/pfu at 38.5°C)
Sabin 2/T79A GCC	0.3	1
S2/2A/79Aa GCG	0.2	1.4
S2/2A/79Ab GCC	1.1	1.8
Sabin 2 ACC	0.1	>2
S2/2A/79Ta ACG	1.2	3.2
S2/2A/79Sa UCG	0.3	1.3
S2/2A/79Sb AGC	0.1	2.7

TABLE 3.9 Table of log drop in titre at 37.5°C and 38.5°C compared to that at 35°C for viruses with different codons but the same amino acids for residue 79. Synonymous amino acids were alanine, threonine and serine.

There would appear to be some effect of RNA sequence in addition to the amino acid effect observed as not all the viruses with the same amino acid have comparable values for drop in titre with temperature. Polioviruses were reported to have no codon preference, unlike other picornaviruses (Stanway *et al.*, 1990). Any differences in ts of viruses with synonymous codons would therefore be unlikely to be due to codon usage unless BGM cells have an unexpected preference. Changing RNA could affect any secondary structure formed. Coding RNA is unlikely to have a highly ordered structure as it would limit evolutionary sequence drift, although from phylogenetic studies 2A has a different evolutionary path from the other picornaviral proteins. It is possible therefore that the RNA for 2A could have a dual purpose of coding for a polyprotein and some regulatory RNA function like part of the hepatitis C IRES. It is conceivable that the viruses have other changes and further sequencing should be carried out to ensure that the 5'NCR has

not reverted in some viruses. In addition problems were encountered with the assay method and limited time did not allow for any repeats to confirm the results presented here.

3.8.2 EFFECTS OF AMINO ACID AT RESIDUE 80 OF 2A

Four viruses with different amino acid changes introduced at residue 80 of 2A were assayed in BGM cells along with a non ts Sabin 2 virus with a G at 481 which was used as a control. As can be seen from table 3.10 the virus with the Sabin 2 codon for phenylalanine at residue 80 was ts. Perhaps not surprisingly the virus with at leucine a residue 80, S2/2A/80L, was relatively non ts since a change of phenylalanine to leucine at this position was found in a non ts revertant of LLΔ472. The LLΔ472 virus is much more ts than Sabin 2, and the 2A change that is able to compensate for a deletion at 472 is also able to compensate for a change of G to A at 481 in Sabin 2. In addition, results shown in table 3.4 above confirm that a single change in 2A is able to compensate for more than one domain V mutation and that there are no direct links between nucleotide changes in domain V and 2A. The S2/2A/80V virus with a valine at residue 80 was also found to be as non ts as S2/2A/80L whereas S2/2A/80I was as ts as a Sabin 2 virus.

VIRUS	Log_{10}(pfu at 35°C/pfu at 38°C)	Log_{10}(pfu at 35°C/pfu at 39°C)
Sabin 2 - 481 G	0.3	0.6
Sabin 2/T79A	0.4	1.0
Sabin 2	0.1	2.0
S2/2A/80V	0.3	0.5
S2/2A/80L	0.2	0.3
S2/2A/80I	0.2	2.2
S2/2A/80F	0.5	2.3

TABLE 3.10 Table of results from a ts assay in BGM cells of Sabin 2 viruses with different amino acids at residue 80. The log drop in titre was calculated at 37.5°C and 38.5°C from that at 35°C. A non ts Sabin 2 virus with a G at 481 was used as comparison along with the original revertant with a change at residue 79.

The results from these experiments indicate that not all amino acid substitutions at this position are able to compensate equally for ts phenotypes resulting from domain V mutations. There does not however appear to be any simple rule about the effect of size or electrodensity of amino acid substitution as indicated above in fig 3.4. In addition there also appears to be some effect of code used for codon 79 but further investigations should be carried out to study this.

3.9 EFFECT OF 2A CHANGES ON MONKEY NEUROVIRULENCE

All the evidence for the translational enhancing effect of the 2A mutations described above was carried out *in vitro* using ts as a measure. Although ts and attenuation have been shown in some cases to be manifestations of the same genotype it was not necessarily true of this phenomenon. A monkey neurovirulence test was carried out to WHO guidelines (WHO 1983), except that fewer animals were used, comparing an attenuated Sabin 2 virus (481A), a virulent Sabin 2 virus with a G at 481 and Sabin 2 viruses carrying either the H96Y or the Y19H 2A mutations.

The results, illustrated in table 3.11 show that the Sabin 2 variants with 2A mutations H96Y and Y19H had attenuated phenotypes comparable to the attenuated phenotype of Sabin 2. Both the clinical score and the mean lesion score recorded for these viruses were similar, contrasting with the results for Sabin 2 with a G at 481. Ts phenotypes of the 2A mutations were however similar to Sabin 2/481G rather than Sabin 2 (Macadam *et al.*, 1994). This would therefore suggest that although 2A appears to have some role in translation, in addition to being responsible for shut off of host cell translation, the effect of this does not change the virulence of the virus for monkeys by intraspinal inoculation. This suggests that the ability of the protease 2A to affect translation is either cell specific or

species specific and any changes in 2A arising during vaccine manufacture or passage in the human gut may have no implications on increasing neurovirulence.

VIRUS	CLINICAL SCORE	MLS (RANGE)	VIRUS PHENOTYPE
SABIN 2	0/4	0.21 (0.14-0.27)	ATTENUATED/TS
SABIN 2/481G	3/4	1.25 (0.93-1.83)	VIRULENT/NON TS
SABIN 2/H96Y	0/4	0.22 (0.12-0.37)	ATTENUATED/NON TS
SABIN 2/Y19H	0/4	0.34 (0.23-0.51)	ATTENUATED/NON TS

TABLE 3.11 Results of the monkey neurovirulence test with Sabin 2, a Sabin 2 with a G at 481 and two Sabin 2 viruses with compensating 2A changes selected in BGM cells. The clinical score refers to the number of animals showing signs of paralysis in one or more limb. MLS refers to the mean lesion score determined by histological analysis of stained spinal cord slices.

The 2A changes discussed here were all selected after passage of virus in BGM tissue culture. Although viruses grow quite well in BGM cells, different selection pressures may be exerted on passage in the human gut. This might suggest that 2A is exerting its effect from interaction with cellular factors whose expression may be cell specific. These results further define the domain V secondary structure as a location of determinants of neurovirulence. Both the attenuated Sabin 2 virus and the non ts revertants with 2A changes had a disruption to domain V secondary structure and were unable to cause paralysis. In contrast the Sabin 2 with a G at 481 and the non-disrupted domain V secondary structure was neurovirulent. What is not apparent from this is how well the non ts revertants with 2A changes would grow in the gut. It is feasible that the enhanced growth of these viruses observed at higher temperatures in BGM tissue culture would also be observed on passage in the gut. Changes in 2A could enable a virus to grow better in the gut and thus able to bring about an improved immune response but with no selection pressure to revert in the 5'NCR. This would make such viruses ideal vaccine candidates.

3.10 DISCUSSION

The work described in this chapter consolidated the observation that amino acid changes in the protease 2A could compensate for disruptions in domain V and render a virus non ts (Macadam *et al.*, 1994). A total of twenty eight Sabin 2 and Leon/Lansing viruses that were non ts on assaying in BGM cells were sequenced through domain V and all were found to have retained the domain V mutations originally responsible for their ts phenotypes. On sequencing through the gene for the protease 2A, viruses were instead found to have one coding change.

Although only three of the 2A coding changes were previously reconstructed into a Sabin 2 virus and confirmed to be able to compensate the ts phenotype (Macadam *et al.*, 1994), the result can presumably be extrapolated to the rest of the changes found. All viruses were found to carry one base mutation and all of them were found to alter amino acids. If the changes were simply the result of evolutionary drift and errors made during RNA replication, silent base changes would have also been found. Indeed some viruses would have been found to have no changes at all in this region. Clearly there appears to be some selection pressure exerted by BGM cells to produce a 2A coding change. Furthermore, viruses with 2A changes were still attenuated when monkeys were intraspinally inoculated in a WHO neurovirulence test. This indicated that the phenomenon was cell specific or simply an *in vitro* observation and that changes in 2A would not alter the neurovirulence of vaccines. Therefore a disruption to domain V was a determinant of attenaution and ts whereas a change in 2A was a determinant of ts only.

Overall, changes were found at twenty two different residues along the length of the 2A polyprotein although most were in the first 100 of the 150 amino acids. A virus with a single domain V mutation could be rendered non ts by more than one 2A change, at least

five, and the same 2A change could be selected in viruses with different domain V mutations. In addition the same 2A mutations could be selected from the same virus on separate occasions. Although there was no trend of type of substitution that occurred, except perhaps the preference of transition over transversion, not all of the changes were equivalent. This was illustrated both by site-directed mutagenesis at positions 79 and 80 and by comparison of the phenotypes arising from the five different 2A changes found in LL472/537 UG derived viruses. Different amino acids at either residues 79 or 80 produced viruses with different ts phenotypes with no link between size, shape or electrodensity. The effect of the changes in 2A also had limitations. Despite the fact that a change at residue 80 was found to compensate a ts phenotype resulting from deletion of nucleotide 472, when the change was introduced into a virus with an even more severe phenotype due to disruption of the entire stem, no effect was observed. In addition, changes introduced into the codon for residue 79 suggested some role for RNA sequence.

The viral protease 2A is known to cleave the initiation factor eIF4G which causes the shut off of host protein synthesis. The protein translation machinery in the cells can then be employed to synthesise only viral proteins in a cap-independent manner. The results presented here show that 2A has another, possibly cell specific, role during protein synthesis. Exactly how the mutations in 2A exert their effect is not known. To have found so many different mutations is perhaps surprising and may suggest that the effect is an artefact. If all the changes had been found to cluster around a very small area then it would be easy to assume that the function is localised on the protein at that area and the amino acid changes are able to improve the function. Alternatively if the changes had been found near the catalytic triad or putative catalytic site the protease activity might be altered. This is not the case and no effect on shut off was observed in Macadam *et al.*, (1994).

However the number of mutations may not be so surprising if the ability of 2A to bind to a cellular factor is considered. The mutations could conceivably disrupt a protein-protein binding function. Therefore by reducing the ability of 2A to bind to a cellular factor, that factor is released to play a part in protein synthesis. Many cellular factors have been found to bind to the IRES and some act to enhance the efficiency of protein synthesis rather than play a essential role. A cell specific factor would have to act in a similar way, perhaps as an RNA 'chaperone' (Svitkin *et al.*, 1996). This would also explain why many viruses were selected with 2A changes rather than domain V reversions. There would be many ways in which a binding function could be destroyed but usually only one or two ways domain V could revert. Therefore a virus is more likely to mutate in 2A and remain viable than in domain V.

An RNA binding function of 2A cannot be completely ruled out but it is unlikely. No known RNA binding motifs can be identified and no evidence for it has been obtained. Despite this, all the changes in 2A were detected in viruses with disruptions to the secondary structure of RNA and there would appear to be some kind of possible link. 2A could somehow be binding to the disrupted domain V region and allowing it to re-form the secondary structure. However, as described above, the large number of mutations throughout the protein suggest that a binding function is disturbed. In addition the possible cell specificity of the enhancing function of 2A suggests interaction with a cell specific protein factor.

If the cell factor binding hypothesis is correct then the cellular factor would be present in BGM cells but not in monkey neuronal cells. This could therefore be tested by looking for differences in cDNA libraries made from the two types of cell. Purified 2A protein could also be used with an expression library to test for binding. Production of recombinant 2A

is unfortunately technically difficult as expression in eukaryotic cells is lethal, causing the shut off of cap-dependent translation (Davies *et al.*, 1991). Placing the 2A gene under HIV tat activation allowed a permanent inducible cell line to be made (Sun & Baltimore 1989). Expression in bacteria does not guarantee correct folding but appears to be sufficient for some activity (Sommergruber *et al.*, 1989; Sommergruber *et al.*, 1992; Liebig *et al.*, 1993).

Recombinant protein production in large quantities would also enable the crystal structure of 2A to be determined. Currently only putative predictions exist from information already obtained about serine proteases. Consequently the positions of the changes found here cannot be accurately plotted although as 2A is a very small protein the positions of changes described here would probably be found on the surface. Effects of changing amino acids could also be calculated using computer programs.

The results presented in this chapter lead on to further investigations. As mentioned above, the cellular factor which is interacting with 2A could be identified. More non ts revertants could be characterised to perhaps identify all the possible 2A changes. It could take a very long time but may give further insight into how the changes affect the protein once the crystal structure is determined. Reversion in other cell lines could identify more cell specificity and present more scope to the identification of the interacting cellular factor.

Further study should be carried out as to how a Sabin strain virus with a change in 2A behaves as a vaccine. It is possible that the attenuating change at 481 (type 2 numbering) would be genetically stable in culture and *in vivo* when in combination with a change in 2A and make a safer vaccine. A study of the evolution of virus in the monkey CNS or in a chimpanzee feeding experiment would indicate just how stable the virus is. At this point in time the WHO is particularly concerned with complete eradication of poliomyelitis in the

very near future and perhaps a change in vaccine at this late stage is unlikely. The attenuated vaccine is under much debate in this initiative as it poses a problem of excretion of live virus in vaccinees. However any information gained could be useful in the future development of other enterovirus vaccines.

More work should be carried out on the viruses with changes introduced at residue 79. It is unclear why the viruses with different codons but the same amino acids should have such different ts phenotypes. A new regulatory role for the coding RNA is possible and cannot be completely outruled. Codon usage may also be more important than previously calculated in either polioviruses or the eukaryotic cells in which the virus polyprotein is translated. Alternatively other changes could have been introduced in the clones during their construction and sequencing of the 5'NCR should be carried out first. This experiment could also be carried out using an alternative residue to confirm whether changes in RNA are significant. Residues could also be altered in positions where changes have not yet been found to determine how much the protein can be changed. In addition the natural variation of 2A in polioviruses or enteroviruses as a whole could be determined.

CHAPTER FOUR

COMPARISONS OF WILD TYPE SEQUENCES

4.1 INTRODUCTION

The previous chapter described the discovery of twenty two coding changes in the poliovirus protein 2A that are able to rescue a ts phenotype resulting from disruption to the secondary structure of the 5'NCR. Heterogeneity of 2A at the amino acid level was measured at less than 7% from the consensus as a result of comparison of only seven naturally occurring viruses, including the three Sabin strains, in Macadam *et al.*, (1994) and reproduced here in fig 4.1. If this is a true representation of variation then the selection of such a large range of changes described in the previous chapter is perhaps surprising.

```
                1                                                   50
P1/Sabin        ---------- ---------- ---------- -n-------- -t--r-----
P1/Mahoney      ---------- ---------- ----d----- -n---s---- -t--r-----
P2/Sabin        ---------- ---------- ---------- ---------- -a--r-l---
P2/Lansing      ---------- ---------- ---------- -ni--i---- -v------i-
P3/Leon         ---------- ---------- ---k-----t --i------- -v--------
P3/Sabin        ---------- ---------- ---k------ --i------- -v--------
P3/Finland      ---------- ---------- ---------- ---------- -t------i-
Consensus       GFGHQNKAVY TAGYKICNYH LATQEDLQNA VSVMWNRDLL V-ESKAQGTD

                51                                                  100
P1/Sabin        ---------- ---------- ---------- ------n--- ----------
P1/Mahoney      ---------- ---------- ---------- ------n--- ----------
P2/Sabin        -----s--t- ---------- -----i---- ------e--- ----------
P2/Lansing      -------ht- ---------- -----t---- ------e--- ----------
P3/Leon         ---------- ---------- ---------- ---------- ----------
P3/Sabin        ---------- ---------- ---------- ---------- ----------
P3/Finland      -------st- -------sr- ---------- ---------- ----------
Consensus       SIARCNCNAG VYYCESRRKY YPVSFVGPTF QYMEANDYYP ARYQSHMLIG

                101                                                 149
P1/Sabin        ---------- -----h---- ---------- ---t------ ---------
P1/Mahoney      ---------- -----h---- ---------- ---------- ---------
P2/Sabin        ---------- ---------- ---------- ---------- ---------
P2/Lansing      ---------- ---------- ---------- ---------- ---------
P3/Leon         ---------- ---------- --v------- ---------- ---------
P3/Sabin        ---------- ---------- --v------- ---------- ---------
P3/Finland      ---------- ---------- ---------- ---------- ---------
Consensus       HGFASPGDCG GILRCQHGVI GIITAGGEGL VAFSDIRDLY AYEEEAMEQ
```

FIGURE 4.1 Amino acid sequences of poliovirus 2A . Sequences were from the GenBank/EMBL database and amino acid sequence obtained and analysed using the 'translate', 'pileup' and 'pretty' programs on GCG. The threshold was set at 1.5 when using 'pretty' so that all amino acid differences were displayed. Conserved residues are shown as dashes (-) and differences to the consensus are shown as lower case letters. From Macadam *et al.*, (1994).

The first complete RNA sequence of an entire poliovirus was published in 1981. Since that time many other complete sequences have been published and can be found on the GenBank/EMBL Database. Although this contains both complete and partial sequence data of wild type and Sabin virus strains the number of wild type sequences that have been elucidated remains small. The database holds several partial sequences of 2A but these contain information of only the first 60 bases (Rico-Hesse *et al.*, 1987). The story is much the same for the 5'NCR of the virus although many more partial sequences of only the 5'NCR exist in the database.

The objective of the work described here was to obtain more sequence data of wild type viruses for comparison and to determine how much the protein 2A has varied over time. Evidence from genetic and phylogenetic comparisons of viruses from the enterovirus genus suggests that the capsid and the 3'NCR regions seem to have evolved together, indicating a link between the two (Pöyry *et al.*, 1996). The 5'NCR has undergone little evolution, probably due to importance of conservation of secondary structure and 2A has evolved by itself, possibly because it acts alone (Pöyry *et al.*, 1996). More sequence data on domain V of the 5'NCR was also desirable in light of the results of the previous chapter.

RESULTS

4.2 WILD TYPE VIRUSES

The National Institute for Biological Standards and Control, where this project was based, holds many samples of serotyped poliovirus extracted from clinical samples. These have been collected from patients with various forms of the disease poliomyelitis for the last five decades and from all over the world. Viruses are grown in tissue culture either at the Institute or country of origin and stored at -70°C in tissue culture fluid. Nine such viruses

were selected from this stock for sequencing, three of each serotype, dating from 1952 to 1991 and isolated from six different countries (see table 4.1).

SEROTYPE	NAME IN TEXT	ISOLATE	YEAR OF ISOLATION	PLACE OF ISOLATION
TYPE 1	P1.1	Lennette 77675	1955	Hawaii
	P1.2	Morocco 5/78	1978	Morocco
	P1.3	Gam 92	1986	Gambia
TYPE 2	P2.1	Lennette 77728	1953	California, USA
	P2.2	Morocco 92/78	1978	Morocco
	P2.3	Pakistan 2/90II	1990	Pakistan
TYPE 3	P3.1	Lennette 77730	1952	California, USA
	P3.2	Morocco 109/76	1976	Morocco
	P3.3	India 40/91	1991	India

TABLE 4.1 List of wild type viruses selected for sequencing. All nine were taken from stocks held at the National Institute for Biological Standards and Control.

4.2.1 SEQUENCING THE 2A GENE

Viruses were initially grown in HEp-2C monolayers in 25cm^2 flasks and RNA was extracted from tissue culture fluid using phenol/chloroform. This was then used to synthesise cDNA for template in a PCR reaction. Primers used previously on the non ts revertants in chapter three, AM20 and AM21 (see chapter two) were initially tested with an annealing temperature of 37°C so as to allow hybridisation despite the expected mismatches. This however caused problems with contamination and non specific binding. Raising the annealing temperature in later cycles of a PCR reaction to encourage specific binding produced no improvement and using the higher temperature for all cycles resulted in no product at all. Consequently new primers were designed with degenerate bases from sequence information already available on the database. These primers, 3'WTAR and ARWTI (see chapter two) were again tried together and with the corresponding pair from AM20 and AM21 but again the lower annealing temperature produced contamination problems and the higher temperatures produced no product. Specifically making cDNA

with the antisense primer ARWTIII instead of the random hexamers routinely used had no effect, again probably because of the large number of mismatches.

It was then decided that direct RNA sequencing may produce better results as shorter primers are used in these reactions. These would have less chance of encountering problems with primer hybridisation. Viruses were again grown in 25cm^2 flasks of HEp-2C monolayers and purified on a 15-30% sucrose gradient. Pelleted RNA was then used in RNA sequencing reactions with primers All13, All14½ (see chapter two) and ARWTIII. The first two were designed to be complementary to Sabin strains of virus and had previously been successfully used on wild type viruses (personal communication, G. Dunn). Unfortunately no readable sequence could be obtained using this method which was probably due either to the inefficiency of primer hybridisation or to poor quality of RNA.

In an attempt to again make a PCR product from these viruses which could then be sequenced more primers were designed with fewer degenerate bases from comparison with coxsackieviruses and polioviruses from the GenBank/EMBL database. Coxsackieviruses show the highest homology to polioviruses when comparing nucleotide sequence similarity of the P1 region (Palmenberg 1989), the 5'NCR and to a smaller extent the P2 region (Pöyry *et al.*, 1996). These primers, WTPmlI3266 and WTBstE3939 (see chapter two) incorporate the restriction sites PmlI and BstE respectively for possible future cloning of PCR products. A PCR reaction of 30 cycles was carried out where the first 5 had an annealing temperature of 37°C and this was raised to 50°C for the next 25. This meant that in the early rounds of amplification the primers could bind with a little less specificity than the later rounds when the majority of the templates would be identical to the primer. From

this, PCR products were obtained without contamination which were cleaned using the Prep-a-Gene silica matrix method.

Purified DNA was then used as template in a T7 sequencing reaction. Unfortunately only the WTBstE3939 primer was able to produce any sequence but from information this provided, two other primers were designed, WT#3 and WT#4 (see chapter two), from which more of the 2A gene sequence was obtained. Unfortunately not all of the 2A coding sequence was acquired and the sequencing is incomplete. The sequences of the wild type viruses obtained were compared and the result of the 'pileup' and 'pretty' programs in the Genetics Computer Group are displayed in appendix 1. Sequences were then translated using the 'translate' program and the comparison is displayed in fig 4.2.

The nucleotide sequence comparison displays a high third base variation that would indeed present primer hybridisation problems. The comparison of the amino acid sequences however shows that the variation is much less and that most of the nucleotide variation produces silent changes. Both the PCR and sequencing primers would appear to be more specific to type 2 viruses as the majority of the 2A sequence was actually obtained for these viruses. Despite this, not all the changes appear to be specific to a serotype although complete variance does occur at many positions. In addition there is no evidence of evolution over time in any of the serotypes from either the amino acid or nucleotide sequences. All the viruses would appear to be the result of individual outbreaks as none of these viruses seem to be direct descendants from the earlier sample.

```
                1                                                    50
P1.1            .......... .......... .......... .......... ..........
P1.2            .......... .......... .......... .......... ..........
P1.3            .......... .......... .......... .......... ..........
P2.1            .......... ---s------ --k---p--- ---------- i----sp---
P2.2            .......... .......... .-----sa-- --v------- ----------
P2.3            .......... .......... .......... .......... .........-
P3.1            .......... .......... .......... .......... ..........
P3.2            .......... .......... .......... .......... ..........
P3.3            .......... .......... .......... .......... ..........
P3/LEON         gfghqnkavy ---y------ ---k--l--t ---------- -v--------
CONSENSUS       ---------- TAG-KICNYH LATQED-QNA VSIMWNRDLL VTESKAQGTD

                51                                                   100
P1.1            .......... .......... ........-- ---------- ----------
P1.2            .......... .....y---- ---------- ---------- --c-------
P1.3            .......... ..--d----- ---------- ------e--- ----------
P2.1            ---------- ---------- ---------- ----v----- ----------
P2.2            --------a- ---------- ---------- ------e--- ----------
P2.3            -----s---- ---------- ---------- ------e--- ----------
P3.1            .......... .......... .....g---- ---------- ---------a
P3.2            .......... .......... .......... .......... ..........
P3.3            .........- ---------- ----yv---- ---------- ----------
P3/LEON         --------a- ---------- -----v---- ---------- ----------
CONSENSUS       SIARCNCNTG VYYCESRRKY YPVSFIGPTF QYMEANDYYP ARYQSHMLIG

                101                                                  149
P1.1            --........ ..f------- ---------- ---------- ---------
P1.2            ---------- ---------- ---------- -------... .........
P1.3            ----a----- ---------- --v------- ---------- ---------
P2.1            ---------- ---------- ---------- --l....... .........
P2.2            ----xx-e-- --------x- ---------- ---------- ---------
P2.3            ---------- ---------- ---------- ---------- ---------
P3.1            y--------- ---------- --v------- -----l.... ....---q-
P3.2            .......... .--------- --v------- ---------x x--------
P3.3            ---------- ---------- ---------- d---n---x- ---------
P3/LEON         ---------- ---------- --v------- ---------- ---------
CONSENSUS       HGFASPGDCG GILRCQHGVI GIITAGGEGL VAFSDIRDLY AYEEEAMEQ
```

FIGURE 4.2 A comparison of the 2A amino acid sequences from the nine wild type viruses sequenced. Leon is also included so as to orientate the incomplete sequence data. The amino acid sequence was calculated from the nucleotide sequence using the GCG 'translate' program and the alignment was made using 'pileup' and 'pretty' programs with the threshold set at 1.5. Conserved residues are shown as dashes (-) and differences to the consensus are in lower case. Blanks are shown as full stops (.) where no sequence was obtained and the x denotes cross bands and the places where sequence could not be read.

4.3 COMPLETE VIRUS SEQUENCES FROM DATABASE

All complete poliovirus sequences that are available on the GenBank/EMBL database were selected for comparison with sequence data obtained here and details of them are listed in

table 4.2. However it must also be noted that the correct Sabin 2 sequence was later published by Pollard *et al.*, (1989) but as there are no differences in the regions under study here the Toyoda *et al.*, (1984) sequence can be directly taken. Mahoney, Leon/37, Leon 119 and Finland viruses were all isolated from cases of poliomyelitis whereas Lansing and the W2 virus were mouse adapted viruses, Lansing being virulent, W2 being attenuated. Although Leon 119 virus was isolated as a virulent revertant of Sabin 3 the 2A and domain V sequences of these were found to be identical and therefore Leon 119 is omitted from amino acid comparisons. Virus W2 is the only addition to the fig 4.1 amino acid comparison from Macadam *et al.*, (1994).

STRAIN	SEROTYPE	ACCESSION NUMBER	REFERENCE
Sabin 1	I	V01150	Nomoto *et al.*, 1982
Sabin 2	II	X00595	Toyoda *et al.*, 1984/ Pollard *et al.*, 1989
Sabin 3/Leon a1b	III	X00596	Toyoda *et al.*, 1984/ Stanway *et al.*, 1983
Mahoney	I	J02281	Kitamura *et al.*, 1981/ Racaniello & Baltimore 1981a
Lansing	II	M12197	LaMonica *et al.*, 1986
W2	II	D00625	Pevear *et al.*, 1990
Leon 119	III	X01076	Cann *et al.*, 1984
Leon/37	III	K01392	Stanway *et al.*, 1984a
Finland	III	X04468	Hughes *et al.*, 1986

TABLE 4.2 List of viruses whose complete sequences are available on the GenBank/EMBL database and used in comparison of wild type viruses sequenced here in the 2A gene and domain V of the 5'NCR. Accession numbers and original references are listed.

A comparison of nucleotide sequence can be found in appendix 2. Leon 119 is included to show that the sequence is the same as for Leon. As for the figure in appendix 1 for the wild type viruses sequenced here the extent of third base variance throughout the entire

region can easily be seen. A similar comparison of 2A amino acid sequences to that displayed in fig 4.1 but including W2 is displayed in fig 4.3. The same GCG computer programs (GCG Programme Manual 1994) were used for its construction. The addition of W2 changes the consensus slightly at amino acids 33 and 42 and W2 has two coding changes at places not previously listed and the consensus at position 33. Otherwise the figure given for amino acid variation in Macadam *et al.*, (1994) is not altered.

```
                1                                                   50
P1/Sabin        ---------- ---------- ---------- -nv------- -t--r-----
P1/Mahoney      ---------- ---------- ----d----- -nv--s---- -t--r-----
P2/Sabin        ---------- ---------- ---------- --v------- -a--r-l---
P2/Lansing      ---------- ---------- ---------- -ni--i---- --------i-
P2/W2           ---------- ---------- ---------- ini--i---- --------i-
P3/Leon         ---------- ---------- ---k-----t --i------- ----------
P3/Sabin        ---------- ---------- ---k------ --i------- ----------
P3/Finland      ---------- ---------- ---------- --v------- -t------i-
Consensus       GFGHQNKAVY TAGYKICNYH LATQEDLQNA VS-MWNRDLL VVESKAQGTD

                51                                                  100
P1/Sabin        ---------- ---------- ---------- ------n--- ----------
P1/Mahoney      ---------- ---------- ---------- ------n--- ----------
P2/Sabin        -----s--t- ---------- -----i---- ------e--- ----------
P2/Lansing      -------ht- ---------- -----t---- ------e--- ----------
P2/W2           -------ht- ---------- ---------- ------e--- ----------
P3/Leon         ---------- ---------- ---------- ---------- ----------
P3/Sabin        ---------- ---------- ---------- ---------- ----------
P3/Finland      -------st- -------sr- ---------- ---------- ----------
Consensus       SIARCNCNAG VYYCESRRKY YPVSFVGPTF QYMEANDYYP ARYQSHMLIG

                101                                                 149
P1/Sabin        ---------- -----h---- ---------- ---t------ ---------
P1/Mahoney      ---------- -----h---- ---------- ---------- ---------
P2/Sabin        ---------- ---------- ---------- ---------- ---------
P2/Lansing      ---------- ---------- ---------- ---------- ---------
P2/W2           ---------- ---------- ---------- ---------- ---v-----
P3/Leon         ---------- ---------- --v------- ---------- ---------
P3/Sabin        ---------- ---------- --v------- ---------- ---------
P3/Finland      ---------- ---------- ---------- ---------- ---------
Consensus       HGFASPGDCG GILRCQHGVI GIITAGGEGL VAFSDIRDLY AYEEEAMEQ
```

FIGURE 4.3 2A amino acid sequences of all published complete poliovirus. Sequences were from the GenBank/EMBL database. Amino acid sequences were obtained and analysed using the 'translate', 'pileup' and 'pretty' programs on Genetics Computer Group. The threshold was set at 1.5 when 'pretty' was used so as to display all amino acid differences. Conserved residues are shown as dashes (-) and differences to the consensus are shown as lower case letters.

The first third amino acids show the most variation from the consensus with eleven out of fifty positions varying and the last third show highest homology with only four out of forty nine positions varying. This is in contrast to the nucleotide sequence which shows the same variation throughout the region. Therefore more of the third base changes in the last third must be silent than in the first third. This suggests that the last third of the amino acids make up an area of the 2A protein that has a more fundamental role in the functioning of the protease, perhaps near the active site. Until the crystal structure of the protein has been deciphered it is impossible to decide where all the residues actually lie.

A complete comparison of all the wild type viruses taken from the database and all the information obtained from the nine wild type viruses sequenced here is shown in figure 4.4. The relatively high conservation of amino acid sequence can be seen with some evidence that changes are serotype specific. Some variation of the viruses sequenced here was found at positions where changes were found in the sequences from the database *e.g.* 14, 23, 27, 28, etc., and some variation was found at new positions *e.g.* 33, 47, 56, 59 etc. Although not all the sequences are complete the first fifty amino acids of the protease would be predicted to have the highest variation.

```
                1                                                  50
P1/Sabin        ---------- ---------- ---------- -nv------- ----r-----
P1/Mahoney      ---------- ---------- ----d----- -nv--s---- ----r-----
P1.1            .......... .......... .......... .......... ..........
P1.2            .......... .......... .......... .......... ..........
P1.3            .......... .......... .......... .......... ..........
P2/Sabin        ---------- ---------- ---------- --v------- -a--r-l---
P2/Lansing      ---------- ---------- ---------- -ni--i---- -v------i-
P2/W2           ---------- ---------- ---------- ini--i---- -v------i-
P2.1            .......... ---s------ --k---p--- --i------- i----sp---
P2.2            .......... .......... .-----sa-- --v------- ----------
P2.3            .......... .......... .......... .......... .........-
P3/Leon         ---------- ---------- ---k-----t --i------- -v--------
P3/Sabin        ---------- ---------- ---k------ --i------- -v--------
P3/Finland      ---------- ---------- ---------- --v------- --------i-
P3.1            .......... .......... .......... .......... ..........
P3.2            .......... .......... .......... .......... ..........
P3.3            .......... .......... .......... .......... ..........
CONSENSUS       GFGHQNKAVY TAGYKICNYH LATQEDLQNA VS-MWNRDLL VTESKAQGTD
```

```
                  51                                                  100
P1/Sabin          --------a- ---------- ---------- ------n--- ----------
P1/Mahoney        --------a- ---------- ---------- ------n--- v---------
P1.1              .......... .......... .........- ---------- ----------
P1.2              .......... .....y---- -----i---- ---------- ----------
P1.3              .......... ..--d----- -----i---- ------e--- ----------
P2/Sabin          -----s---- ---------- -----i---- ------e--- ----------
P2/Lansing        -------h-- ---------- -----t---- ------e--- ----------
P2/W2             -------h-- ---------- ---------- ------e--- ----------
P2.1              ---------- ---------- -----i---- ----v----- ----------
P2.2              --------a- ---------- -----i---- ------e--- ----------
P2.3              -----s---- ---------- -----i---- ------e--- ----------
P3/Leon           --------a- ---------- ---------- ---------- ----------
P3/Sabin          --------a- ---------- ---------- ---------- ----------
P3/Finland        -------s-- -------sr- ---------- ---------- ----------
P3.1              .......... .......... .....g---- ---------- ---------a
P3.2              .......... .......... .......... .......... ..........
P3.3              .........- ---------- ----y----- ---------- ----------
CONSENSUS         SIARCNCNTG VYYCESRRKY YPVSFVGPTF QYMEANDYYP ARYQSHMLIG

                  101                                                 149
P1/Sabin          ---------- -----h---- ---------- ---t------ ---------
P1/Mahoney        ---------- -----h---- ---------- ---------- ---------
P1.1              --........ ..f------- ---------- ---------- ---------
P1.2              ---------- ---------- ---------- -------... ..........
P1.3              ----a----- ---------- --v------- ---------- ---------
P2/Sabin          ---------- ---------- ---------- ---------- ---------
P2/Lansing        ---------- ---------- ---------- ---------- ---------
P2/W2             ---------- ---------- ---------- ---------- ---v-----
P2.1              ---------- ---------- ---------- --l....... ..........
P2.2              ----xx-e-- --------x- ---------- ---------- ---------
P2.3              ---------- ---------- ---------- ---------- ---------
P3/Leon           ---------- ---------- --v------- ---------- ---------
P3/Sabin          ---------- ---------- --v------- ---------- ---------
P3/Finland        ---------- ---------- ---------- ---------- ---------
P3.1              y--------- ---------- --v------- -----l.... ....---q-
P3.2              .......... .--------- --v------- ---------x x--------
P3.3              ---------- ---------- ---------- d---n---x- ---------
CONSENSUS         HGFASPGDCG GILRCQHGVI GIITAGGEGL VAFSDIRDLY AYEEEAMEQ
```

FIGURE 4.4 Full comparison of 2A sequences including data from the eight viruses from the database and the nine wild type viruses sequenced here. The figure was complied using 'pileup' and 'pretty' GCG programs with the threshold set at 1.5. Conserved bases are shown as dashes (-) and differences to the consensus in lower case. Missing sequence is shown as full stops (.) and positions where the sequence was not determined denoted as (x). Positions of compensating 2A changes are indicated in red in the consensus.

4.4 CONSERVATION OF 2A IN THE ENTEROVIRUS GENUS

Sequence comparison of poliovirus 2A above highlighted areas of conservation and areas of variation. To further analyse the evolution of this protease, sequences from other

picornaviruses were also compared. Viruses were selected from only complete sequences, on the GenBank/EMBL database and accessed using GCG.

As depicted in chapter one (fig 1.2) polioviruses are members of the enterovirus genus along with coxsackieviruses and enteroviruses. Other close phylogenetic relatives, determined by sequence analysis, are rhinoviruses and to a lesser extent, apthoviruses (Palmenberg 1989). The 2A protease of apthoviruses is much shorter and the leader protein is instead responsible for cleavage of eIF4G resulting in shut off of host cell synthesis. It was therefore decided to include only sequences from coxsackieviruses, enteroviruses and rhinoviruses. Echoviruses were previously all classified as enteroviruses and because of the close sequence similarity to coxsackieviruses one echovirus sequence was also included along with a sequence of a swine vesicular virus. All the viruses listed below in table 4.3 also have similar IRES and 5'NCR regions.

NAME	ACCESSION NUMBER	REFERENCE
HUMAN RHINOVIRUS 1B	D00239	Hughes *et al.*, 1988
HUMAN RHINOVIRUS 2	X02316	Skern *et al.*, 1985
HUMAN RHINOVIRUS 14	K02121	Stanway *et al.*, 1984b Callahan *et al.*, 1985
HUMAN RHINOVIRUS 16	L24917	Lee *et al.*, 1995
HUMAN RHINOVIRUS 89	M16248	Duechler *et al.*, 1987
COXSACKIEVIRUS A9	D00627	Chang *et al.*, 1989
COXSACKIEVIRUS A16	U05876	Pöyry *et al.*, 1994
COXSACKIEVIRUS A21	D00538	Hughes *et al.*, 1989
COXSACKIEVIRUS A24	D90457	Supanaranond *et al.*, 1992
COXSACKIEVIRUS B1	M16560	Iizuka *et al.*, 1987
COXSACKIEVIRUS B3	M33854	Lindberg *et al.*, 1987 Klump *et al.*, 1990
COXSACKIEVIRUS B4	X05690	Jenkins *et al.*, 1987
COXSACKIEVIRUS B5	X67706	Zhang *et al.*, 1993
ECHOVIRUS 11	X80059	Auvinen & Hyypiä 1990
ECHOVIRUS 12	X79047	Kraus *et al.*, 1995
SWINE VESICULAR DISEASE VIRUS H/3 '76	X54521	Inoue *et al.*, 1989
ENTEROVIRUS 70	D00820	Ryan *et al.*, 1990
BOVINE ENTEROVIRUS VG-5-27	D00214	Earle *et al.*, 1988

TABLE 4.3 Table of all viruses included in a comparison of 2A sequences from viruses in the enterovirus genus. Viruses are listed along with their accession number on the GenBank/EMBL database with the journal reference.

The 2A sequences obtained from these viruses were translated into the amino acid sequence using the 'translate' GCG program. A comparison of all the sequences, including those obtained for polioviruses was calculated using 'pileup' and 'pretty' GCG programs. The results obtained are reproduced in fig 4.5. As only complete sequences are used the sequences for the viruses listed in table 4.1 and sequenced here were not included.

```
                1                                                  50
Rhino 14        .-l-pry-gi -ts-v--m-- --m-p--hh- liapyp---- ai.vstggh-
Rhino 1B        ...-psdly- h---liyr-l --.fnsemhd sil-s-ss-- ii-r-n-.t-
Rhino 2         .........- h---liyr-l --.fnsemhe sil-s-ss-- ii-r-n-.v-
Rhino 16        .tv-psdmy- h---liyr-l --.fns-ihd sil-s-ss-- ii-r---.q-
Rhino 89        .........- h---liyr-l --.fns--dd sil-s-ss-- ii-r-n-fe-
Lansing         .------k-- -tag------ ---------- --ni-wi--- --.ve-k-q-
W2              .------k-- -tag------ ---------- -ini-wi--- --.ve-k-q-
Sabin 1         .------k-- -tag------ ---------- --n--w---- --.te-r-q-
Mahoney         .------k-- -tag------ -----d---- --n--ws--- --.te-r-q-
Sabin 3         .------k-- -tag------ ----k----- --si-w---- --.ve-k-q-
Leon            .------k-- -tag------ ----k----- --si-w---- --.ve-k-q-
Sabin 2         .------k-- -tag------ ---------- --s--w---- --.ae-r-l-
Finland         .------k-- -tag------ ---------- --s--w---- --.te-k-q-
Cox A24         -------m-- --ag------ ----p--hd- --r-lw---- mi.vs-r-q-
Cox A21         .------k-- --ag------ ----ps-hl- -is-lwd--- m-.ve-r-q-
Cox A9          -a--q-s--- -----rvi-r ----ht-w-- c--ed----- --.s-t--h-
Cox A16         -k--q-s--i -----rvv-r ----hn-wa- l--e-ss--- --.-st--q-
Cox B1          -a--q-s--- -----rvv-r ----r--w-r c--ed----- --.s-t--h-
Cox B5          -vl-q-t--i ci---rvv-r ----s--w-r c--ed----- --.s-t--h-
Cox B4          ...---s--- ------vv-r ----hv-w-- c--ed----- --.s-t--h-
Cox B3          ...-q-s--- -----rvv-r ----sa-w-- c--es----- --.s-t--h-
Echo 11         ...-y-s--- --v--rvv-r ----ht-w-- c--ed----- -i.s-t--h-
Echo 12         ......... ......... .........- c--e------ --.--t--h-
CoxB4           ...---s--- ------vv-r ----hv-w-- c--ed----- --.s-t--h-
SVD             -a--q-s--- -----rvv-i ----ra-w-- c--ed-d--- --.s-t--h-
Entero VG       -p--q-q--a ---s---l-r ----ya-we- e--qs-q--- --.trvd-h-
Entero 70       .......-ga f--s---i-- ----d-eker s-y-dwqs-v --.t-va-h-
CONSENSUS       GGFGHQNGAV YVGNYKICNY HLATQED-QN AVWVMYNRDL LVYSTSTA-G

                51                                                 100
Rhino 14        aet-ph---- s----styyr ----iic-k- -niwiegnp- --s-f-agvm
Rhino 1B        d-y-ps---- eat---khkn r---ikvtph dwy-iq-s-- --khi-ynl-
Rhino 2         d-y-ps-d-- qat---khkn r-f-itvtsh dwy-iq-s-- --khi-ynl-
Rhino 16        dgy-pt---- eat---khkn r---invtph dwy-iq-s-- --khi-.nl-
Rhino 89        n-v-pn-d-- ect---hhkd r-f-irvtah dwy-iq-s-- --khi-ynl-
Lansing         i--------h ---------r -------t-- ---y-ean-- --a-------
W2              i--------h ---------r -------v-- ---y-ean-- --a-------
Sabin 1         t--------n a--------r -------v-- ---y-eann- --a-------
Mahoney         t--------n a--------r -------v-- ---y-eann- --v-------
Sabin 3         t--------n a--------r -------v-- ---y-eand- --a-------
Leon            t--------n a--------r -------v-- ---y-eand- --a-------
Sabin 2         t------s-n ---------r -------i-- ---y-ean-- --a-------
Finland         i--------s ---------s r------v-- ---y-eand- --a-------
Cox A24         s-t------r ------k-mk -----tvte- ---y-eand- --a---t---
Cox A21         t-t----s-r c--------r ---l-t-t-- --rf-rand- --a-------
Cox A9          --v----g-- ----f-a-kn -h-------- glv-vq-s-- --k-----v-
Cox A16         -------d-q ------s--r -h-----sk- slifvea--- --------lm
Cox B1          --i----q-- ----f-a--n -h-------- glv-vq-s-- --k-----v-
Cox B5          --t----r-s ----f-a--n -h-------- glv-vq-s-- --k-----v-
Cox B4          --t----q-- ----f-a-ks -h-------- clv-vq-s-- --k-----v-
Cox B3          --i----q-- ----f-a-kn -h--i----- glv-vq-s-- --r-----v-
Echo 11         --v----r-s ------q-kg -h---n---- glv-vq-s-- --k-----v-
Echo 12         -------q-- ----f-a--n -h-------- .lv-vg---- --r-----v-
SVD             --t----d-- a---f-a--n -h---t---- glv-vq-s-- --kkh---v-
Entero VG       --t------r s-i---k-ta -h--ivvtp- siykieand- --e-m-t-i-
Entero 70       khq----r-n ------khkn rs---c---- gi-win-sd- --a---tnt-
CONSENSUS       CDSIARCNCT TGVYYCESR- KYYPVSFEGP TFQEM-ESEY YPARYQSHML
```

```
            101                                                    151
Rhino 14    k-c-p----- -------i-- p--ll----s -y-s---u-q -ecia--q...
Rhino 1B    --e-pc---- ---k-l-r-- ----i ----- -h---t-l-q fqca--q....
Rhino 2     --e-pc---- ---k-l-k-- -----t---d nh---i-l-h fhca--q....
Rhino 16    --e-pc---- ---k-l-k-- ----i----- -h---i-l-h fhca-......
Rhino 89    --e-pc---- ---k-l-k-- ---mi----- -h---i-l-k fqca--q....
Lansing     ------s--- -------q-- ----i ----- -----s---- -yay-------
W2          ------s--- -------q-- ----i----- -----s---- -yay-v-----
Sabin 1     ------s--- -------h-- ----i----- -----t---- -yay-------
Mahoney     ------s--- -------h-- ----i----- -----s---- -yay-------
Sabin 3     ------s--- -------q-- ---------- -----s---- -yay-------
Leon        ------s--- -------q-- ---------- -----s---- -yay-------
Sabin 2     ------s--- -------q-- ----i----- -----s---- -yay-------
Finland     ------s--- -------q-- ----i----- -----s---- -yay-------
Cox A24     l-m------- -------n-- -m-------n -i-------- -wvy-------
Cox A21     --c------- -------t-- ----i----r -I-------- -wvy-------
Cox A9      laa--s---- -------e-- ------m--- -v-g---v-- ---l-dd----
Cox A16     lav-hs---- -------q-- -v---st--n ---g---v-- ---ld------
Cox B1      laa--s---- -------e-- -v----m--- -v-g---va- ---l-dd----
Cox B5      laa--s---- -------e-- ---l--m--- -v-g---v-- ---l-dd----
Cox B4      lat--sf--- -------e-- ---l--m--- -v-g---v-- ---l-dd----
Cox B3      laa--s---- -------e-- ------m--- -v-g------ ---l-dd----
Echo 11     laa--s---- -------e-- ------m--- -v-g---v-- ---l-dd----
Echo 12     laa--s---- -------e-- ---l--m--- -v-g---v-- ---l-dd----
SVD         laa------- -------q-- ------m--- -v-g---v-- ---l-dd----
Entero VG   l-i------- ---l---e-- -m--l-v--g dh-g---v-- ---i-dd----
Entero 70   lam-pcq--- ---l-v-s-- ---l------ -i---t---n ---l-dd----
CONSENSUS   IGHGFAEPGD CGGILRC-HG VIGIVTAGGE GLVAFADIRD LLW-EEEAMEQ
```

FIGURE 4.5 A comparison of 2A amino acid sequences from rhinoviruses, polioviruses, coxsackieviruses (Cox) and enteroviruses (Entero). Sequences were obtained from the GenBank/EMBL database. Amino acid sequence was obtained using the 'translate' program on GCG. The comparison was compiled using 'pileup' and 'pretty' GCG programs. The threshold was set at 1.5 when using 'pretty'. Differences to the consensus are shown in lower case letters and positions of compensating 2A changes found in non ts revertant polioviruses and discussed in chapter three are indicated in red. Other abbreviations are echoviruses (Echo) and swine vesicular disease virus (SVD).

An immediate observation from the comparison above is that although there are small areas of high conservation across all of the viruses listed, there are three clear groups: the rhinoviruses; the polioviruses (including coxsackievirus A21 and coxsackievirus A24) and the coxsackieviruses which also includes enteroviruses, echoviruses and the swine vesicular disease virus. The coxsackieviruses and the polioviruses also show more sequence homology to each other than to the rhinoviruses. These results are of course

completely expected from past phylogenetic studies. Numbering of residues is slightly different as the proteins are of slightly different lengths with some residues either missing or inserted within the different groups.

The putative catalytic triad residues of 2A, numbered here as 21, 39 and 111 are completely conserved. The residues found to be highly conserved in small serine proteases, numbered here as 40, 90, 91 and 126 are mostly conserved except for one change at residue 40 in enterovirus 70 and two changes at 126, in rhinovirus 2 and in coxsackievirus A16. This is also an expected result as all 2A sequences are able to act as a protease. In addition, the highest variation is found in the first two thirds of the amino acid sequence with more conservation found in the last third. This was also found when comparing only polioviruses and reflected in the 2A compensating mutations also being found mainly in the first two thirds. High conservation of the region made up of the forty nine residues implies it is important to the function of the whole protein, possibly forming part of the active site or playing a role during folding. Indeed the largest run of conserved amino acids throughout all the viruses is the five residues at 109-113 which are GDCGG. This must be a very important region to the protease and appears to be found very close to the active site on the three dimensional model predicted by Bazan & Fletterick (1988) and Yu & Lloyd (1991) and displayed in chapter one (fig 1.11).

Positions of compensating 2A changes found in non ts revertants of polioviruses discussed in chapter three are shown in red in the consensus sequence in figure 4.5. Within the polioviruses, positions of the changes are generally highly conserved but when other viruses are included the picture is a little more complicated. Only two of the changes, at residues 50 and 108, are in positions of total conservation (106 in poliovirus numbering). Most of the changes are at positions that are conserved within the three groups but differ

between groups. In addition some of the changes are in positions where polioviruses share conservation with rhinoviruses and some where polioviruses share conservation with coxsackieviruses.

These results give some insight about the compensating changes. Only two changes were found at a position of complete conservation which suggests that the fundamental protease activity that all the proteins share is not affected. Most changes were found to be at positions that were conserved only within a group. This would perhaps suggest that the changes are altering a poliovirus specific function of 2A. Therefore no effect would be expected from the introduction of equivalent changes into viruses outside of the poliovirus group.

4.5 DOMAIN V OF THE 5'NCR

Results from the previous chapter presented some link between the 5'NCR disruptions of domain V and the gene 2A. Therefore it was decided that comparison of domain V sequence data from the nine wild type viruses selected from NIBSC stocks and listed in table 4.1 may show the same genetic drift as the 2A sequence data. A sequence comparison of the wild type viruses obtained from the GenBank/EMBL database is displayed in fig 4.6 and was constructed in the same way as for the 2A sequences.

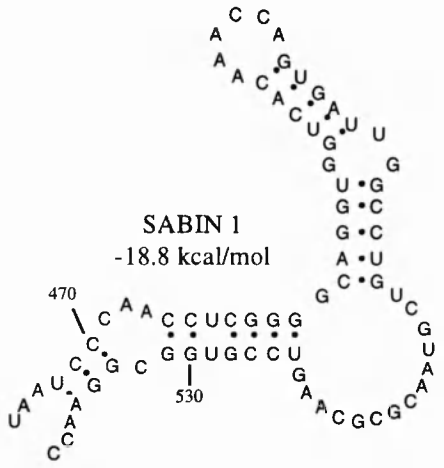
```
                464                                                 513
P1/Sabin        ---------- c----u---g -----ug--- a--------u --uug-----
P1/Mahoney      ---------- c----u---- -----ug--- a--------u --uug-----
P2/Sabin        ---------- ---------- a-----g--- --g------u ---ug--u--
P2/Lansing      ---------- ---------- ---------g ----u----- ----------
P2/W2           ---------- ---------- ---------g ----u----- ----------
P3/Leon         ---------- ------u--- --------cu ----c----- ag--------
P3/Sabin        --------u- ------u--- --------cu ----c----- ag--------
P3/Finland      ---------u c--------- -----u---- a--------u --uuu-----
Consensus       CGGCUAAUCC UAACCACGGA GCAGGCAGUC GCAAACCAGC GACCAGCCUG

                514                         542
P1/Sabin        ---------- ---------- ---------
P1/Mahoney      ---------- ---------- ---------
P2/Sabin        ---------- ------u--- ---------
P2/Lansing      ---------- -----u---- ---------
P2/W2           ---------- -----u---- ---------
P3/Leon         ---------- ---------- ---------
P3/Sabin        ---------- ---------- ---------
P3/Finland      ---------- a--c--t--- ---------
Consensus       UCGUAACGCG CAAGUCCGUG GCGGAACCG
```

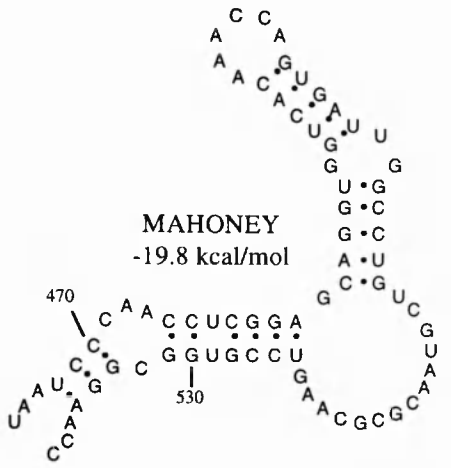
FIGURE 4.6 Domain V sequences of all published polioviruses. Sequences were from the GenBank/EMBL database and analysed using the 'pileup' and 'pretty' programs on Genetics Computer Group. Conserved residues are shown as dashes (-) and differences to the consensus are shown as lower case letters.

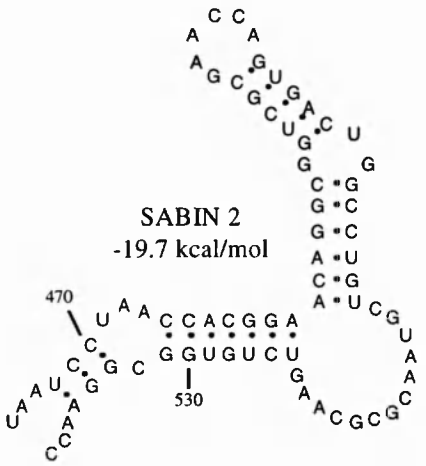
Sequences were then folded into secondary structure using 'foldRNA' and 'squiggles' programs on GCG which also calculate the predicted free energy of the structures. These can be seen in fig 4.7. The structure is most stable in terms of thermodynamics at the lowest free energy value. Therefore when more than one structure can form from one sequence, it is the one with the lowest free energy value that is more likely to form. Other alternatives may also be present in the population at relatively low proportions, particularly when the energy differences are very small and they may exert some effect on the phenotype of the virus.

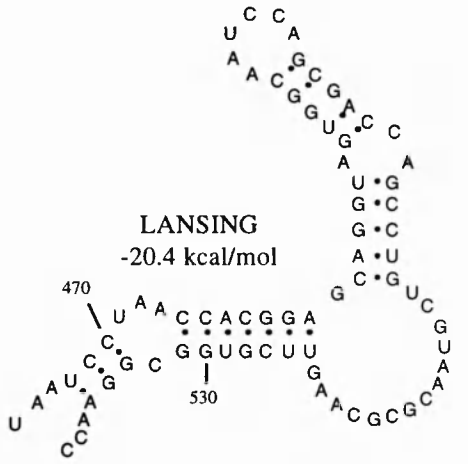
For all domain V structures base 484 and the large loop formed by bases 514-527 were forced to be unpaired following the structure of Skinner *et al.*, (1989) and Pilipenko *et al.*, (1989) published from thermodynamic folding and double and single stranded specific reagents rather than the structure of Le & Zuker (1990). For all published sequences

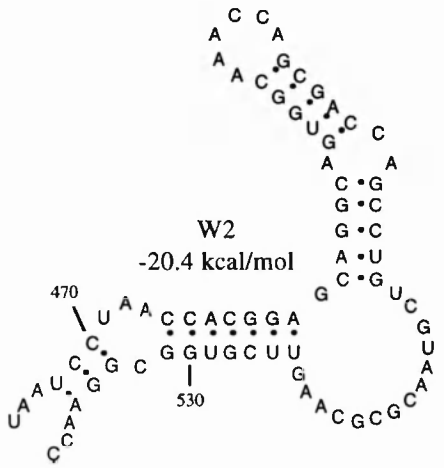
shown here free energy was increased by approximately 3 kcal/mol when this loop was unpaired. Evidence from the Sabin 2 attenuating mutation at 481 suggests that this base is not paired with the U at 511. A change of G to A found in Sabin 2 would make the base pair stronger and it is this mutation which has been shown to attenuate the virus (Ren *et al.*, 1991; Macadam *et al.*, 1991b; Macadam *et al.*, 1993). Sabin 2 is always found to revert back at this position to remove the AU base pair. The loop formed by bases 514 to 527 is highly conserved across both the poliovirus family and the enterovirus genus but there is little evidence to suggest any base pairing from strand specific reagents. An unpaired loop could interact within the 5'NCR or with cellular or viral proteins. In addition, the domain V structures of the viruses from the database do not show the unnamed three base stem made up of bases 464-476 and 580-582 (type 3 numbering) although this was used for free energy calculations.

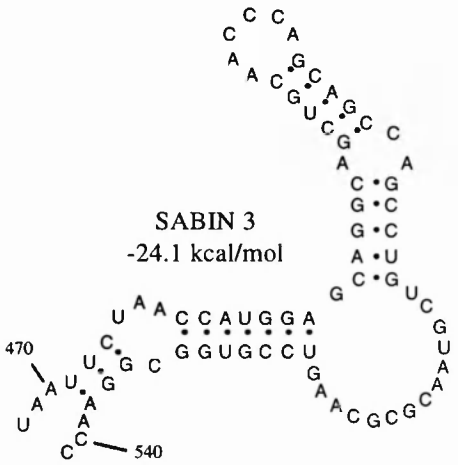

SABIN 1
-18.8 kcal/mol
470
530


MAHONEY
-19.8 kcal/mol
470
530


SABIN 2
-19.7 kcal/mol
470
530


LANSING
-20.4 kcal/mol
470
530


W2
-20.4 kcal/mol
470
530


SABIN 3
-24.1 kcal/mol
470
540

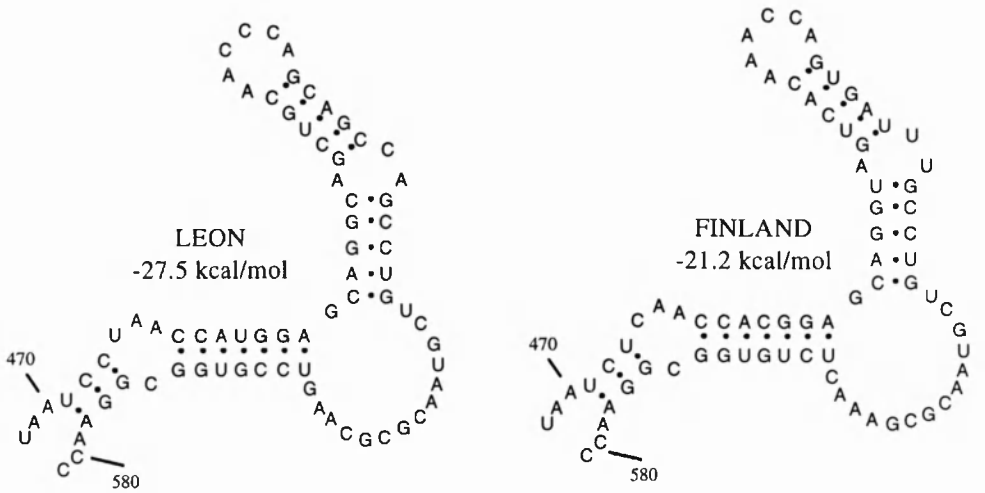


FIGURE 4.7 Domain V structures of the viruses whose entire sequences are published on GenBank/EMBL database. Structures and free energy values were determined using 'foldRNA' and 'squiggles' programs.

Free energy calculated from predictions of secondary structure can be correlated with the temperature at which a virus displays a ts phenotype for viruses that only differ in domain V (Macadam *et al.*, 1992). Values presented in fig 4.7 show that there is no strict correlation with free energy and virulence. Mahoney has a relatively high free energy in comparison to the values for Sabin 1 and Sabin 2. In contrast Sabin 3 has a relatively low free energy value, lower than the values for Lansing and W2. Clearly free energy values are a valuable tool to determine which of the alternative structures are more likely to form when the sequence could allow variations. However these values are a guide and do not necessarily reflect the complex interactions, both within the rest of the RNA structure and with protein factors, that are fundamental to the lifecycle of the virus.

4.6 SEQUENCING DOMAIN V OF THE 5'NCR

Domain V proved easier than 2A to sequence as the primers PCR F and PCR 9 used previously with the non ts revertants worked well on the wild type viruses. PCR products were purified using the Prep-a-Gene matrix method and sequenced with primers 13/II and αLL390 (see chapter two). However sequence was obtained from only eight of the viruses

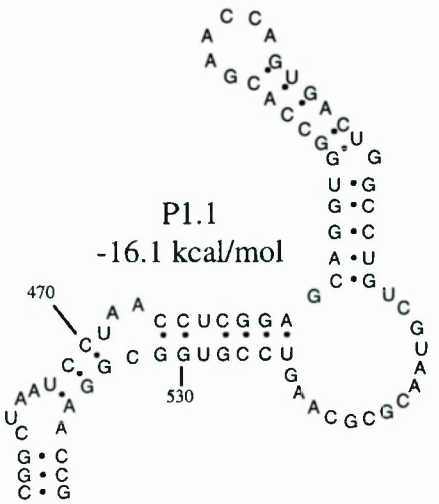
as no PCR product could be made for the P3.3 virus. Fig 4.8 displays the domain V sequence comparison between these viruses calculated from 'pileup' and 'pretty' GCG programs.

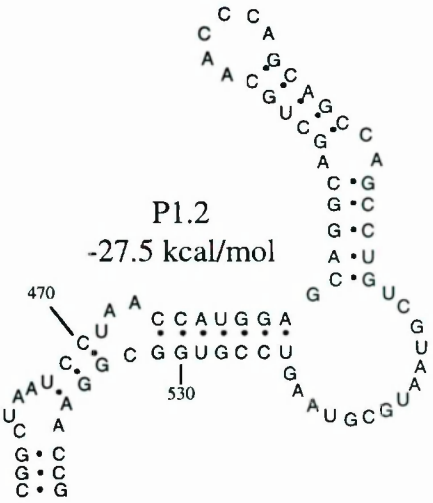
```
             464                                                  513
P1.1         ---------- -----u---- -----u--c- --g------- ----g-----
P1.2         ---------- ------u--- -----ca-cu g---c----c ag-c------
P1.3         ---------- ------u--- -----ca-cu g---c----c ag-c------
P2.1         ---------- ------u--- -----c---- -ug------- -------u--
P2.2         ---------- ---------- -----u-a-- -u--u---a- -gu-g-----
P2.3         ---------- ---------- a----c---- --g------- -------u--
P3.1         ---------u c----u---- -----u-a-- ----g----- --u-g-----
P3.2         ---------- c--------- -----ua--u g--------c a--cu-----
Consensus    CGGCUAAUCC UAACCACGGA GCAGG-GGUC ACAAACCAGU GACUAGCCUG

             514                          542
P1.1         ---------- ---------- ---------
P1.2         ------u--- u--------- ---------
P1.3         ---------- ---------- ---------
P2.1         ---------- ---------- ---------
P2.2         ------g--- u-----u--- ---------
P2.3         ---------- ------u--- ---------
P3.1         ------u--- --------a- aga------
P3.2         ------u--- ------u--- ---------
Consensus    UCGUAACGCG CAAGUCCGUG GCGGAACCG
```

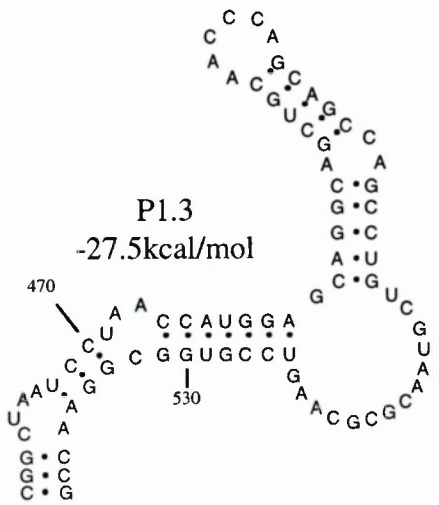
FIGURE 4.8 Comparisons of the 5'NCR domain V sequences of eight of the nine previously un-sequenced wild type viruses. Sequences were compiled using 'pileup' and 'pretty' GCG programs. A consensus sequence is displayed and differences to this are shown as (-).

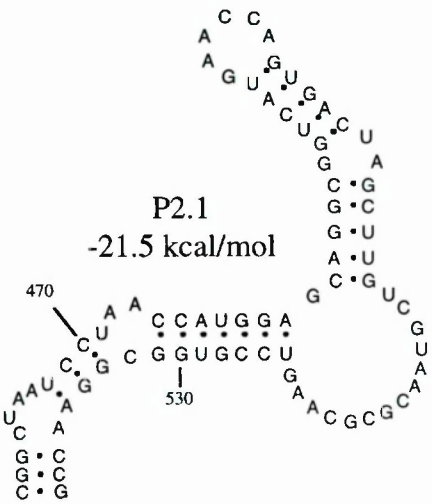
The sequences of domain V were folded and analysed using 'foldRNA' and 'squiggles' GCG programs and can be seen displayed in fig 4.9. The calculated free energy values are also displayed. Many of the viruses have the potential to make an extra GU base pair between bases 490 and either 507 or 508. If this is allowed to occur, the predicted free energy is decreased, making it a more energetically stable structure. However the overall structure of domain V would be altered, increasing either stem (c) or (d) by one base pair and making the 'hinge' at this position more constrained. In the P3.2 virus, this extra base pair is AU and much more likely to form than a GU pair. Other structures and shapes are

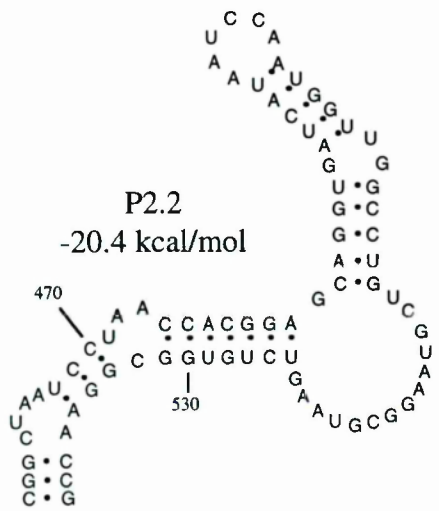
successfully conserved and only the bases around the loop between stems (b) and (c) are forced to pair or unpair in the 'foldRNA' GCG program as discussed above.

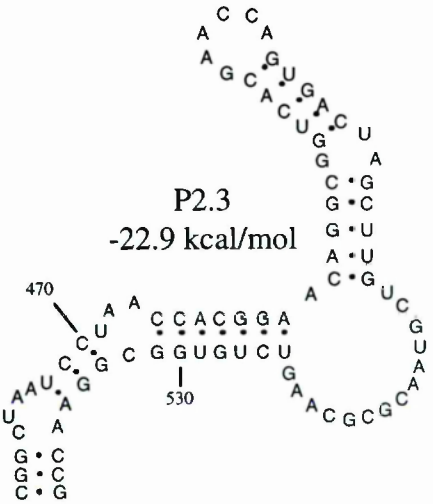












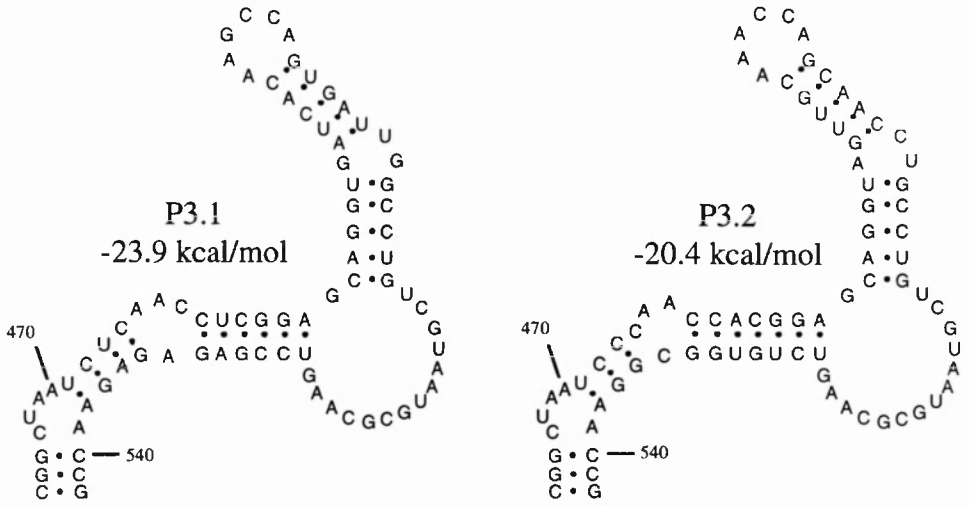


FIGURE 4.9 Predicted domain V secondary structures and calculated free energy values of the eight wild type viruses sequenced here. Structures were determined using 'foldRNA' and 'squiggles GCG programs.

From observation of these structures the virus P2.3 would appear to be vaccine related as it has the Sabin 2 change of G to A at 481. Other changes in the genome must have also occurred to change the neurovirulence. The virus P3.1 appears to have undergone quite a rearrangement at stem (a). The loop between stems (a) and (b) is however conserved as three bases whilst stem (b) is reduced to six base pairs. P1.1 also has different base pairs in stem (d) which appear to reduce the free energy of the structure quite considerably as other stems and loops seem comparable to those found in P1.2 and P1.3.

4.7 SEQUENCE CONSERVATION IN DOMAIN V

When the eight sequences taken from the database are added to the second eight that were sequenced, a clear pattern of sequence conservation is found. Figure 4.10 shows this comparison obtained using 'pileup' and 'pretty' GCG programs. The individual stems and loops of domain V formed by the sequence are also indicated to give a better picture of areas of conservation.

```
                464                                                  502

                STEM LOOP STEM LOOP  STEM B      STEM C    STEM     LOOP
                           A                                  D
P1/Sabin        --- ---- --- c-- --u---g - ----u g ---a- ------
P1/Mahoney      --- ---- --- c-- --u---- - ----u g ---a- ------
P1.1            --- ---- --- --- --u---- - ----u g -c-a- g-----
P1.2            --- ---- --- --- ---u--- - ----- a -cug- --c---
P1.3            --- ---- --- --- ---u--- - ----- a -cug- --c---
P2/Sabin        --- ---- --- --- ------- a ----- g ---g- g-----
P2/Lansing      --- ---- --- --- ------- - ----- a --gg- --u---
P2/W2           --- ---- --- --- ------- - ----- a --gg- --u---
P2.1            --- ---- --- --- ---u--- - ----- g ---au g-----
P2.2            --- ---- --- --- ------- - ----u g a--au --u---
P2.3            --- ---- --- --- ------- a ----- g ---a- g-----
P3/Leon         --- ---- --- --- ---u--- - ----- a -cug- --c---
P3/Sabin        --- ---- -u- --- ---u--- - ----- a -cug- --c---
P3/Finland      --- ---- --u c-- ------- - ----u a ---a- ------
P3.1            --- ---- --u c-- --u---- - ----u g a--a- --g---
P3.2            --- ---- --- c-- ------- - ----u a --ug- ------
Consensus       CGG CUAA UCC UAA CCACGGA G CAGGC - GUC-C AAACCA

                503                                                  542

                STEM      STEM C      LOOP          STEM B      STEM   STEM
                  D                                                  A
P1/Sabin        ----u -g ----- ------------- ------- - --- - ---
P1/Mahoney      ----u -g ----- ------------- ------- - --- - ---
P1.1            ----- -g ----- ------------- ------- - --- - ---
P1.2            -cag- c- ----- ------u---u--- ------- - --- - ---
P1.3            -cag- c- ----- ------------- ------- - --- - ---
P2/Sabin        ----- -g --u-- ------------- --u---- - --- - ---
P2/Lansing      -c--- c- ----- ------------- -u----- - --- - ---
P2/W2           -c--- c- ----- ------------- -u----- - --- - ---
P2.1            ----- -- --u-- ------------- ------- - --- - ---
P2.2            a--gu -g ----- ------g---u--- --u---- - --- - ---
P2.3            ----- -- --u-- ------------- --u---- - --- - ---
P3/Leon         -cag- c- ----- ------------- ------- - --- - ---
P3/Sabin        -cag- c- ----- ------------- ------- - --- - ---
P3/Finland      ----u -u ----- ----------a--c --u---- - --- - ---
P3.1            ----u -g ----- ------u------- ----a-a g a-- - ---
P3.2            -ca-- cu ----- ------u------- --u---- - --- - --a
Consensus       GUGAC UA GCCUG UCGUAACGCGCAAG UCCGUGG C GGA A CCG
```

FIGURE 4.10 A comparison of all the domain V sequences sequenced here and obtained from the GenBank/EMBL database. The figure was compiled using 'pileup' and 'pretty' GCG programs. Conserved bases are shown as dashes (-) and differences as lower case letters. Individual stems and loops that the sequence forms are indicated.

Using figure 4.10 a distinct pattern of conserved bases and non conserved bases can be observed. The areas that show a high level of variation are stem (d) and the short loops either side of it although secondary structure is maintained by covariance. In contrast stem

(a), the adjacent loops and the stem just before it are highly conserved. The loop at the 'top' of domain V, adjacent to stem (d) shows more conservation in the second half (bases 499-501) than the first half. These results compare favourably with comparisons of poliovirus IRES structures in Pöyry *et al.*, (1992) and Jackson *et al.*, (1994) except for P3.1.

Stem sequences could conceivably be regarded as more resistant to genetic drift as both sides of the stem need to mutate together to avoid any disruptions to the secondary structure. It may be interesting to note therefore that loops can be as conserved as stems. This suggests that loops do not merely exist as spacers between stems and their sequence is just as important as for stems. This also suggests an interaction either with other loops of the 5'NCR secondary structure with protein factors involved in replication or translation. In the case of the 'top' loop adjacent to stem (d) the second, more conserved half may play a part in interactions rather than the first half. Alternatively the sequence of a loop may play a part in the stacking of secondary structure *e.g.* the fourteen base loop between stems (b) and (c) is extremely conserved.

4.8 DISCUSSION

The high conservation of 2A at the amino acid level calculated from seven sequences in Macadam *et al.*, (1994) was further corroborated by the addition of sequences from W2 and Leon/119 although a few more residues were found to vary. The comparison of coding RNA sequence for 2A displays the high third base variance that occurs within the different strains and caused problems with accurate primer hybridising. Changes are however not generally specific to each serotype. The sequencing carried out for this project is incomplete but the information obtained further supports the fact that variation of amino acid sequence in 2A is minimal. This observation gives credence to the fact that 2A

changes discussed in chapter three are functional and did not just occur from evolutionary variation.

Comparisons with other picornaviral sequences revealed that for 2A, polioviruses are very similar to coxsackieviruses, particularly coxsackievirus A21 and coxsackievirus A24. This is in complete agreement with previous phylogenetic studies. Conservation was generally quite high within the individual poliovirus, coxsackievirus, rhinovirus and enterovirus groups but considerably less within the whole group. Analysis of the positions of the compensatory changes showed that the protease activity was almost certainly not affected as only two were found at positions of complete conservation and that the effect is probably restricted to polioviruses. In addition a five base GDCGG motif at residues 109-123 which is probably part of the enzymatic active site was also found to be highly completely conserved. The last third amino acids in the sequence in fact displayed the highest conservation, indicating an important fundamental role for this part of the protein.

Conservation of domain V sequence is somewhat dominated by the conservation of functional secondary structure. Comparison of all the sequences, both those taken from the GenBank/EMBL database and those obtained for this project shows that stems (c) and (d) and the end loop made up of bases 498-502 display the highest variation whilst maintaining the secondary structure. In addition the sequences of unpaired loops were found to be as conserved as stems. Loops do not therefore simply exist to act as spacers between stems. Some loops will almost certainly interact with either protein factors or other areas of the 5'NCR.

Calculations of the free energy values of domain V sequences indicated the thermodynamic stability of the domain V secondary structure in isolation is not necessarily linked to the neurovirulence of the virus. Changes that alter the secondary structure of

domain V are clearly determinants of neurovirulence but a certain amount of instability can be accommodated before the virus is non viable. Conservation of the functional structure would appear to be the major factor. Free energy values are however useful tools in the distinction between alternative structures.

Sequencing data could be further consolidated with the completion of the sequencing of both 2A and domain V. New primers can be designed from information obtained here. Perhaps individual primers for each serotype should be made to avoid the use of too many degenerate bases. The first fifty amino acid residues would be predicted to have the highest variation. In addition only data from complete sequences on the GenBank/EMBL database were used here whereas partial sequences of 2A could also have been added.

CHAPTER FIVE

CELL SPECIFIC ACTIVITY OF THE COMPENSATORY ROLE OF 2A

5.1 INTRODUCTION

The non ts revertants of Sabin 2 and Leon/Lansing mutants discussed in chapter three were all selected in BGM cells which derive from monkey kidney cells. Monkeys can be experimentally infected with poliovirus which results in a disease similar to the human form (Couderc *et al.*, 1989; Hashimoto *et al.*, 1984). An African green monkey kidney cell line was found to encode two slightly different poliovirus receptors, $\alpha 1$ and $\alpha 2$, in two separate loci of the genome (Koike *et al.*, 1992). Both are homologous to the human poliovirus receptor in terms of genomic sequence and function. Expression of the receptor may be tissue specific but the BGM cell line appears to express the $\alpha 1$ form (personal communication, A. J. Macadam).

After the successful cloning and characterisation of the human poliovirus receptor gene (Mendelsohn *et al.*, 1989), transgenic mice were made that expressed the human poliovirus receptor gene (Ren *et al.*, 1990). Independently, mouse cell lines were transformed with the cDNA for the human poliovirus receptor. Cells became susceptible to all three strains of poliovirus infection on expression of the receptor (Mendelsohn *et al.*, 1986). The cell line used in this institute is the L20B cell line, originally derived from mouse connective tissue. L20B cells can be infected by all types of poliovirus and normal CPE occurs although they are slightly less sensitive to infection than HEp-2C cells (Pipkin *et al.*, 1993). Viruses also grow as expected *e.g.* Sabin strains are temperature sensitive and wild type viruses are not. In particular the ts phenotype associated with mutations in domain V of the 5'NCR that is expressed in BGM cells is also seen, albeit at lower temperatures (Macadam *et al.*, 1992). Polioviruses do not infect non transformed mouse cell lines.

It was decided to determine how the non ts revertants carrying 2A mutations would behave in the L20B cell line. All the viruses had been selected using only the BGM (monkey) cell

line and their phenotypes in these cells was examined in chapter three. Evidence from monkey neurovirulence tests showed that these non ts viruses were attenuated and this could indicate that the enhancing effect of 2A is cell specific and involves interactions with cellular factors. It was therefore of interest to determine how the non ts revertants behaved in a different cell line and the transformed mouse line was chosen because of the species difference and the similar effects of 5'NCR mutations in BGM and L20B cells.

RESULTS

5.2 TEMPERATURE SENSITIVITY ASSAYS IN L20B MONOLAYERS

Temperature sensitivity plaque assays were carried out in L20B cells to mirror the previous assays in BGM cells. Many of the non ts revertants of Leon/Lansing derived and Sabin 2 viruses that had compensating 2A changes were assayed along with appropriate controls.

5.2.1 LL472/537 UG

Initially four of the non ts revertants of LL472/537 UG were assayed in L20B cells. For controls, the original ts LL472/537 UG and non ts Leon/Lansing (with a CG at 472/537) viruses were used. The drop in titre with temperature for the LL472/537 UG revertants for both BGM and L20B cells are displayed in table 5.1 and fig 5.1.

VIRUS	Log_{10} (pfu at 35°C/pfu at T°C) for BGM cells				Log_{10}(pfu at 35°C/pfu at T°C) for L20B cells			
	38°C	38.5°C	39°C	39.5°C	37.5°C	38.2°C	38.5°C	39°C
LL472/537 UG	0.5	1.7	2.7	3	0.1	2.3	3.0	3.0
UG/2A E25G	0	0.2	0.4	1	0.7	2	2.6	3.7
UG/2A Y93H	0	0.2	0.4	1	0.3	1.1	2.6	3.9
UG/2A P106S	0	0.3	0.9	0.6	0.5	2.3	2.9	3.6
UG/2A S134T	0	0.3	0.5	0.6	0.1	1.3	2.6	3.5
LL (472/537 CG)	0	0.2	0.4	0.75	0.4	0.2	0.4	1.0

TABLE 5.1 List of titre drop with temperature of LL472/537 UG and its derivatives measured as the ratio of pfu at 35°C to that at 38°C, 38.5°C, 39°C and 39.5°C for the BGM assay and at 37.5°C, 38°C, 38.5°C and 39°C for the L20B assay.

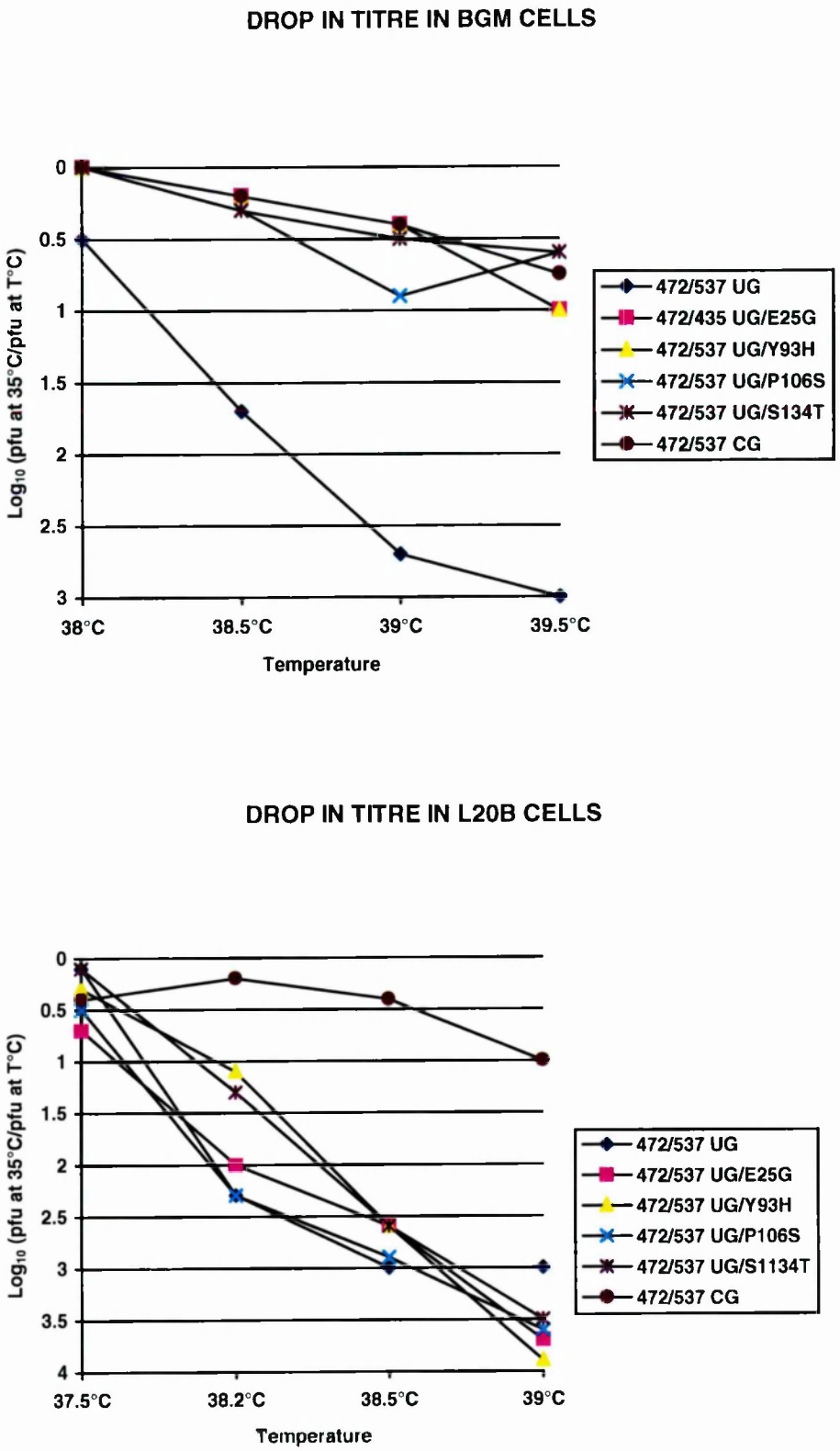


FIGURE 5.1 Graphs to show the drop in titre with temperature for LL472/537UG and its derivatives in both BGM and L20B cell lines.

The results displayed above show that both BGM cells and L20B cells can be used to measure ts phenotypes as LL472/537 UG is ts and the Leon/Lansing (472/537 CG) virus is relatively non ts. But it should also be noted that ts is manifested at a lower temperature in L20B cells and that plaques in L20B cells are smaller (see figs 5.2 and 5.3). The graphs depicting log drop of titre with temperature (see fig 5.1) evidently show that the 2A changes in a LL472/537 UG background had little or no enhancing effect in L20B cells. Clearly all the viruses with 2A changes have similar profiles to the non ts Leon/Lansing in BGM cells and they all have similar profiles to that of the ts LL472/537 UG virus in L20B cells.

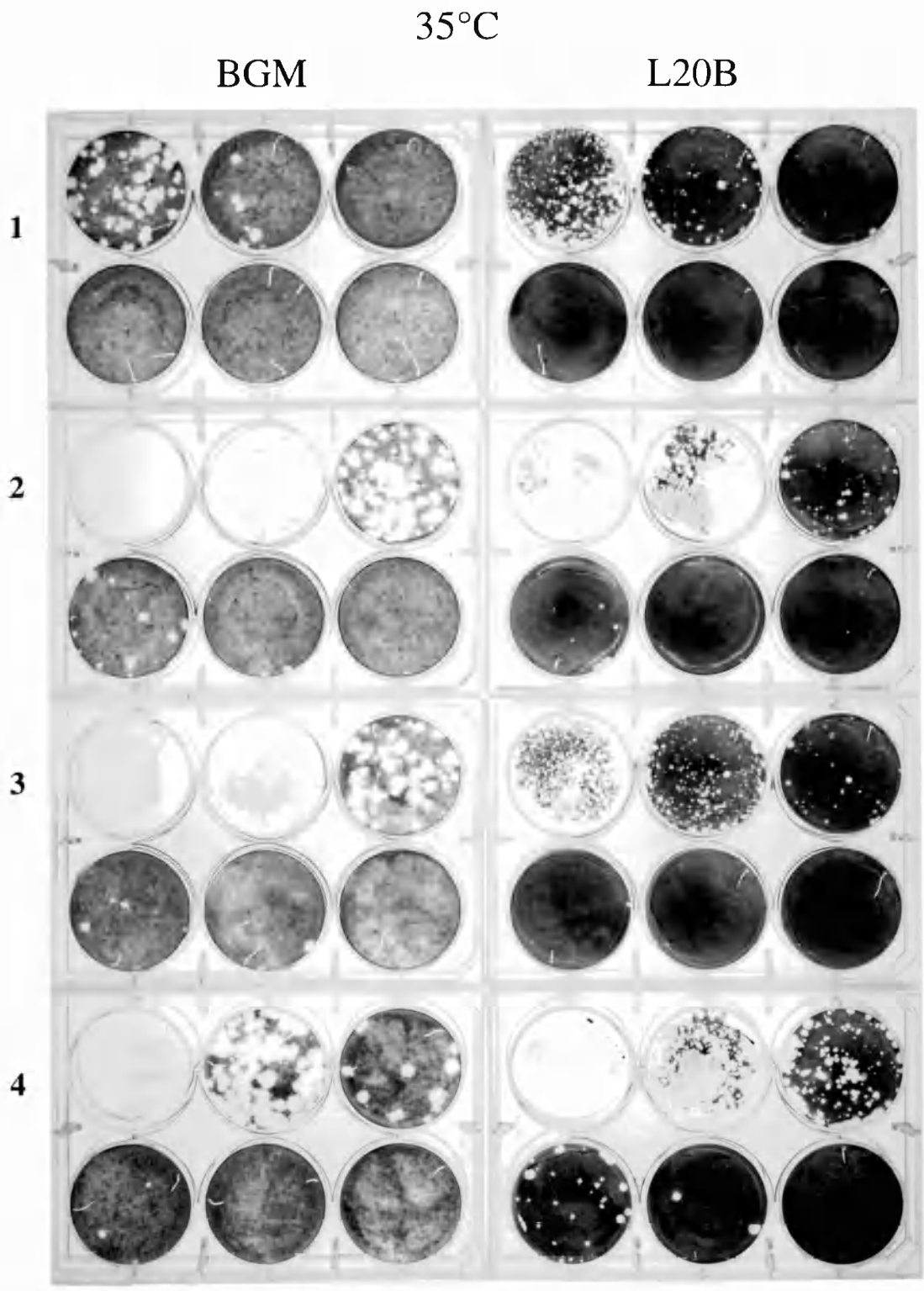


FIGURE 5.2 Plates from a ts assay, on the left in BGM cells, on the right in L20B cells. Inoculated cells were incubated at 35°C with 10 fold dilutions from 10^{-3} to 10^{-8}. The viruses assayed were all Leon/Lansing constructs with a UG mismatch at 472/537. Three viruses also had amino acid substitutions in 2A: 1 = LL472/537 UG; 2 = LL472/537 UG E25G; 3 = LL472/537 UG S134T; 4 = LL472/537 UG P106S.

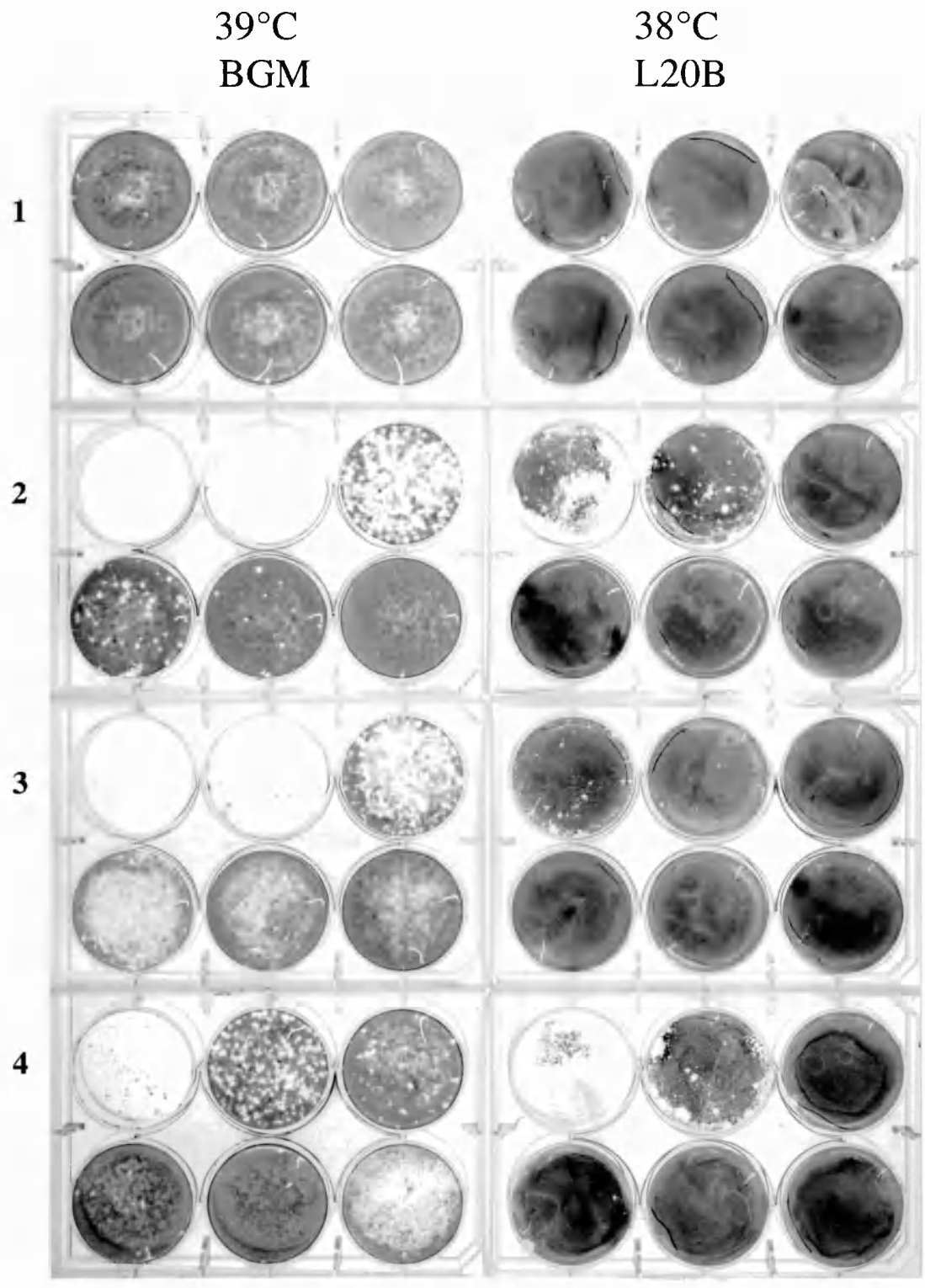


FIGURE 5.3 Plates from a ts assay, on the left in BGM cells, on the right in L20B cells. Cells were inoculated with 10 fold dilutions from 10^{-2} to 10^{-7} and incubated at the relatively higher temperatures of 39°C in BGM cells and at 38°C in L20B cells. The viruses assayed were all Leon/Lansing constructs with a UG mismatch at 472/537. Three viruses also had amino acid substitutions in 2A as indicated: 1 = LL472/537 UG; 2 = LL472/537 UG E25G; 3 = LL472/537 UG S134T; 4 = LL472/537 UG P106S.

A further observation from the graph using data from the L20B cells is that although at the higher temperatures of 38.5°C and 39°C all the non ts revertants have equivalent reductions in titre, at 38.2°C the values are more spread out. This is in contrast to the data from BGM cells where throughout the temperature range all the viruses are more similar to each other. This could mean that although overall the mutations in 2A have little effect on the ts phenotype in L20B cells there could be a limited effect at lower temperatures.

5.2.2 SABIN 2

Similarly, non ts revertants of Sabin 2 were assayed in L20B cells. For comparison, Sabin 2 and the Sabin 2 virus with a G at 481 were included as ts and non ts controls respectively. Again the mutations in 2A appeared to have no enhancing effect as can be seen in table 5.2 and the graphs of fig 5.4. Although the Sabin 2 with a G at 481 is non ts, all the viruses carrying 2A mutations have similar profiles to that of the parent ts virus in L20B cells where as profiles were similar to the non ts Sabin 2/481G in BGM cells. In contrast to the data obtained from LL472/537 UG non ts revertants, those derived from Sabin 2 have much more similar values of drop in titre to each other in both BGM and L20B cells.

VIRUS	Log_{10}(pfu at 35°C/pfu at T°C) for BGM cells				Log_{10}(pfu at 35°C/pfu at T°C) for L20B cells			
	38°C	38.5°C	39°C	39.5°C	37°C	37.5°C	38°C	38.5°C
Sabin 2	0.5	1.25	3.0	4.0	0.75	1.0	1.8	3.25
Sabin 2/481 G	0	0	0.25	1.5	0	0	0.3	0.75
Sabin 2/2A Y19H	0	0	0.25	1.25	0.75	1.0	1.8	2.9
Sabin 2/2A H96Y	0	0	0.25	1.75	0.75	1.0	1.8	3.75
Sabin 2/2A I122V	0	0	0.25	0.75	0.75	1.0	1.8	2.9

TABLE 5.2 List of titre drop with temperature of Sabin 2 and its derivatives measured as the ratio of pfu at 35°C to that at 38°C, 38.5°C, 39°C and 39.5°C for the BGM assay and at 37°C, 37.5°C, 38°C and 38.5°C for the L20B assay.

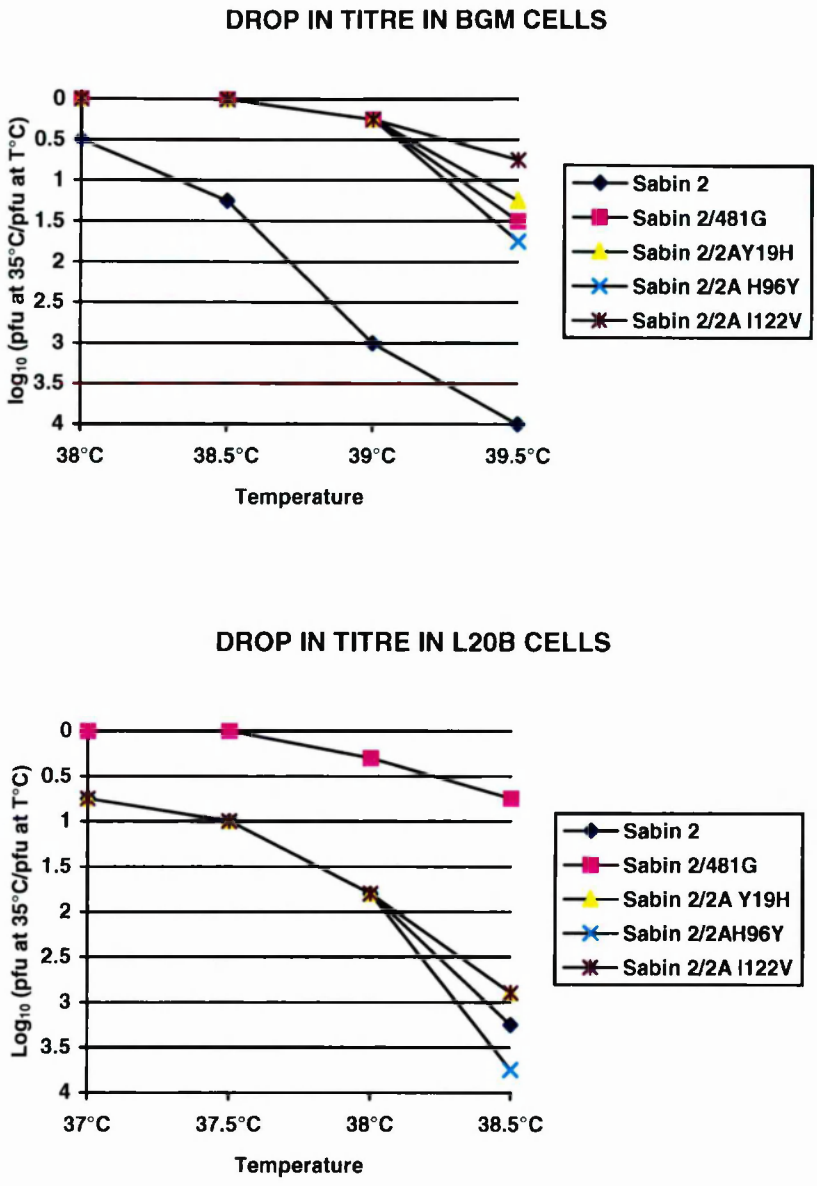


FIGURE 5.4 Graphs to show the drop in titre with temperature for Sabin 2 and its derivatives.

5.2.3 OTHER VIRUSES

Other viruses discussed in chapter three were similarly assayed in L20B cells and the results are displayed in table 5.3. In BGM cells all the viruses with 2A mutations are non ts in comparison with the parent viruses. In contrast the 2A mutations appear to have no suppressive effect on the ts phenotypes in L20B cells. All viruses with 2A mutations were as ts as the parental virus from which they were derived. The results for the LLΔ483

viruses indicate the 2A mutation may have some significance but no consistent and dramatic reduction in ts in obtained in L20B cells like that seen in BGM cells.

VIRUS	Log_{10} (pfu at 35°C/pfu at 39°C) for BGM cells	Log_{10}(pfu at 35°C/pfu at 38°C) for L20B cells
LEON/LANSING	0.1	0.1
LL471/538 AA	1.7	2.1
LL471/538 AA/2A E65K	0.5	2
LL471/538 GA/2A I122V	0.7	1.7
LLΔ472	4.0	>3.4*
LLΔ472/2A F80L	0.5	4.3
LLΔ472/2A C17Y	0.7	4.6
LL473/536 UG	2.0	0.5†
LL473/536 UG/2A K45E	0.5	0.6†
LL473/536 UG/2A T79A	0.4	0.5†
LLΔ483	4.0	5.3
LLΔ483/2A T23I	0.9	3.6
LLΔ483/528	2.5	1.2
LLΔ483/528/2A G48D	0.2	1.7
LL514A	3.0	2.7
LL514A/2A Y10C	0.4	2.8
LL514A/2A Y19C	1.5	2.6
LL514C	1.5	0.5
LL514C/2A E65V	0.2	0.5

TABLE 5.3 Comparison of log drop in pfu in BGM cells and L20B cells for a number of non ts viruses carrying mutations in 2A. For BGM cells it is measured as a ratio of pfu at 39°C and 35°C where as for L20B cells the ratio is pfu at 38°C and 35°C.
* The figure for LLΔ472 is not a definite number as the titre of the virus used was so low that plaques were only obtained at 37°C.
† The figures quoted are from a higher temperature of 39°C as the ts phenotype for LL473/532 UC was not expressed at 38°C.

By comparison with the results for Leon/Lansing the original domain V disruptions resulted in ts viruses in both cell lines whereas the effect of the 2A mutations was restricted to BGM cells. This indicated therefore that the phenomenon was cell specific involving previously unknown interactions with cellular factors. Suppression of ts by 2A is therefore not a simple interaction with domain V.

5.3 NON TS REVERTANTS PICKED IN L20B CELLS

From the results shown above all non ts revertants with 2A mutations plaque picked from BGM monolayers were found to still be ts on assaying in L20B cells. However it was not known whether any other 2A mutations could be selected in L20B cells and function as suppressors of the ts phenotype. It was thus decided to try to select revertant non ts viruses in L20B cells. All structures presented here were predicted using the 'foldRNA' programme of GCG (GCG Program Manual 1994) as described in Macadam et al., (1992).

5.3.1 SABIN 2 REVERTANTS

Cells were infected with Sabin 2 virus in serial log dilutions using a six well plate and incubated at 35°C, 37°C and 38.4°C. After three days plaques were visualised using neutral red stain and five relatively large plaques were picked from cells grown at 38.4°C. Viruses were then grown on HEp-2C monolayers in 25cm^2 flasks and virus was assayed in L20B cells with Sabin 2 and a Sabin 2 virus with a G at 481 for comparison. All plaque derived viruses were found to be similar to Sabin 2 in ts (see table 5.4). Sequencing of the 5'NCR domain V from PCR amplified fragments made using primer PCR F and PCR 9 (see chapter two) and cycle sequencing using a labelled 13/II primer (see chapter two) indicated that all had an A at 481 like Sabin 2.

VIRUS	Log_{10} (pfu at 35°C/pfu at 37.7°C) for L20B cells	Log_{10}(pfu at 35°C/pfu at 38.2°C) for L20B cells
Sabin 2/481G	0.4	0.5
Sabin 2	0.2	1.2
Sabin 2/384.1	0.2	1
Sabin 2/384.2	0.4	1.1
Sabin 2/384.3	0.3	0.8
Sabin 2/384.4	0.3	0.9
Sabin 2/384.5	0.5	1

TABLE 5.4 A comparison of log drop in titre at 37.7°C and 38.2°C as compared to that at 35°C for Sabin 2/484 G, Sabin 2 and the variants plaque picked from it, five were picked at 38.4°C.

These five viruses were further plaque purified at 35°C in L20B cells and two plaques were picked from each. After assaying in L20B cells and sequencing through 481 in domain V it was found that seven viruses were non ts and had a G at 481 and that two viruses were ts and had an A at 481 (see table 5.5). As these viruses had the expected phenotypes from sequence analysis they were not characterised further.

VIRUS	SEQUENCE AT 481	Log_{10} (pfu at 35°C/pfu at 37.7°C) for L20B	Log_{10}(pfu at 35°C/pfu at 38.5°C) for L20B cells
Sabin 2/481 G	G	0.9	1.6
Sabin 2	A	3.1	4.4
Sabin 2/384.1.1	A	2.0	3.1
Sabin 2/384.1.2	G	0.9	1.6
Sabin 2/384.2.1	G	1.0	2.0
Sabin 2/384.2.2	G	0.5	2.3
Sabin 2/384.3.1	A	3.6	4.6
Sabin 2/384.3.2	A	4.0	4.0
Sabin 2/384.4.1	G	1.7	2.8
Sabin 2/384.4.2	G	1.6	2.9
Sabin 2/384.5.1	G	0.5	2.5
Sabin 2/384.5.2	A	3.4	4.5

TABLE 5.5 A comparison of log drop in titre at 37.7°C and 38.5°C as compared to that at 35°C for Sabin 2/484 G, Sabin 2 and the second round of variants plaque picked, two at 35°C from each original plaque derived virus.

No other second site mutation was found in these revertants and this probably reflects the importance of a G at 481. The structural basis of this was further investigated and results are presented in chapter seven. These assays also highlight the inter-assay variation obtained with these cells and the need for control viruses in each assay for comparison of relative effects.

As non ts revertants of Sabin 2 plaque picked in L20B cells were all found to have reverted at 481 in domain V of the 5'NCR it appears that these cells exert little or no selection pressure on the virus to mutate in the protease 2A. This observation is consistent with the observation that the 2A mutations found in revertants selected in BGM cells have no suppressive effect in L20B cells. This suggested that L20B cells would be useful in obtaining further information on the secondary and tertiary structure of the 5'NCR which was the original aim of selecting revertants in BGM cells. If viruses were unable to

compensate for 5'NCR disruptions with changes in the protease 2A they would be forced to compensate for them by changing in the 5'NCR itself or elsewhere. Mutations selected may indicate how other regions interact at these points in domain V or potentially other areas of the 5'NCR. This method would be ideal to select revertants of viruses that cannot directly back mutate. When such viruses were grown in BGM cells non ts revertants all had 2A changes. Of particular interest would be viruses with deletions *e.g.* LLΔ472 or LLΔ483/528.

5.4 NON TS REVERTANTS OF VIRUSES WITH DELETIONS IN THE 5'NCR

Of the many viruses with 5'NCR disruptions, three had base deletions which not only made the virus very ts (see table 3.3) but made direct back mutations extremely unlikely. The viruses of interest were LLΔ472, LLΔ483 and LLΔ483/528 and their predicted domain V secondary structures are depicted below. Non ts revertants of these viruses carrying 2A mutations had already been selected in BGM cells and this procedure was repeated using L20B cells. All structures displayed in this section were constructed from examination of sequences and calculations of folding energies using 'foldRNA' which were displayed using 'squiggles' programmes on GCG. As for structures analysed in chapter four, the large 14 base loop was forced to be unpaired following the structure published in Skinner *et al.*, (1989) and Pilipenko *et al.*, (1989) and the base at 484 (in type 3 numbering) was also forced to be unpaired. Assays were also all carried out in L20B cells to completely remove the possibility of changes in 2A arising during the assay.

5.4.1 REVERTANTS OF LLΔ472

On plating the LLΔ472 virus on L20B cells in a six well plate, nine plaques were selected, five at 35°C and four at 37°C. Viruses from these were assayed and although they were

found to be ts relative to Leon/Lansing, they were not quite as ts as the parent LLΔ472 virus. This can be seen in table 5.6 where the log reduction in titre at the lower temperature of 37.5°C is smaller for the plaque-purified viruses than for the parent virus but all have a similar phenotype at the higher temperature of 38.5°C. On sequencing, all the viruses were found to have a change of G to A at position 534 in domain V in addition to the deletion at 472. No difference was found between the viruses selected at 35°C and those selected at 37°C.

VIRUS	Log_{10} (pfu at 35°C/pfu at 37.5°C) in L20B cells	Log_{10} (pfu at 35°C/pfu at 38.5°C) in L20B cells
Leon/Lansing	0.3	0.6
LLΔ472	2.8	5.0
LLΔ472/35.1	1.4	4.6
LLΔ472/35.2	1.5	4.9
LLΔ472/35.3	1.8	4.8
LLΔ472/35.4	1.4	>3*
LLΔ472/35.5	0.7	>3*
LLΔ472/37.1	1.1	4.5
LLΔ472/37.2	1.7	4.7
LLΔ472/37.3	1.3	>3*
LLΔ472/37.4	1.0	4.1

TABLE 5.6 A comparison of log drop in titre at 37.5°C and 38.4°C as compared to that at 35°C for Leon/Lansing, LLΔ472 and the variants plaque picked from it, five were picked at 35°C and four were picked at 37°C.
* A finite number could not be measured from the dilutions used as the virus titre was so low that plaques were not formed at the higher temperature.

The probable structures of the domain V regions of the parent LLΔ472 virus and the variant virus are shown in fig 5.5. This parent virus has the potential to form two slightly different structures and these are both shown in fig 5.5. The first structure is shown on the left with the three base UAA loop between stems (a) and (b) intact but stem (a) is slightly rearranged. The U at 471 is paired with the G at 537 which originally paired with the C at

472. This results in the now unpaired A at 538 forming a larger loop, leaving stem (a) reduced to two base pairs. The free energy of this structure calculated as -23.2 kcal/mol which is 3.7 kcal/mol higher than that for domain V of Leon. The second version is shown on the right of fig 5.5. The loop between stems (a) and (b) is reduced to AA whilst stem (a) undergoes an alternative rearrangement. The U at 474, previously part of the three base loop, pairs with the G at 530, allowing base pairs 471/537 UG and 473/537 CG to form and reconstruct stem (a) as three base pairs. The free energy of this structure is slightly less at -26.5 kcal/mol, which is only 1.0 kcal/mol higher than that for Leon and the more likely to form thermodynamically. However, the phenotype for LLΔ472 is very ts, indicating that although the domain V structure is stable it is functionally impaired.

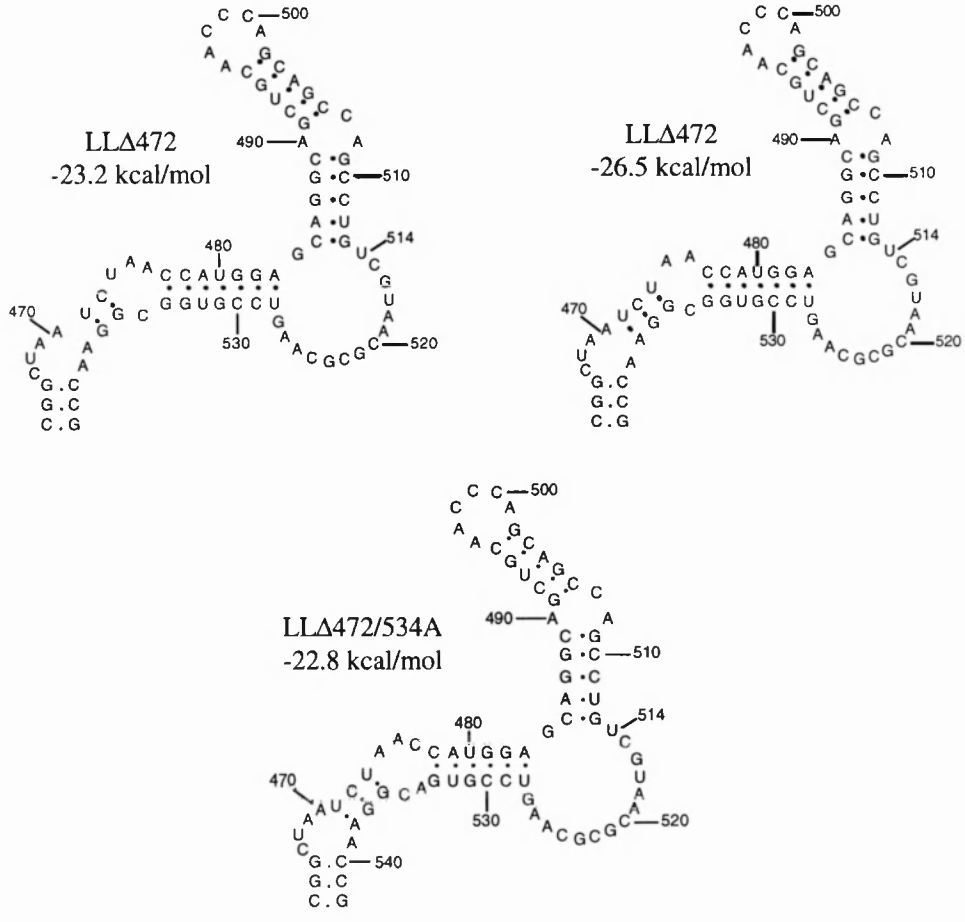


FIGURE 5.5 Possible domain V structures of a LLΔ472 virus and a variant selected from it. Structures of LLΔ472 were calculated in Macadam *et al.*, (1992). Previous type 3 numbering is retained to limit confusion of bases. All structures were analysed using the 'foldRNA' and 'squiggles' programmes of GCG.

The change at 543 in the variant virus disrupts the 477/534 base pair, shortening stem (b) to six base pairs but places extra bases in the loop between stems (a) and (b) (see fig 5.5). Although the loss of a base pair may seem surprising the sequence remaining is such that the loop is able to re-form with three bases as AAC. Stem (a) is also able to re-form as three base pairs whilst the other strand of the loop between stems (a) and (b) is increased. Free energy is calculated as -22.8 kcal/mol (*c.f.* -27.5 kcal/mol for Leon).

The structure predicted here for the LLΔ472 variant would therefore indicate that length or flexibility of the loop between stems (a) and (b) is important to the functional structure of domain V. Whilst the variant is less ts than the parent, it is still ts and has a relatively high free energy value. The change at 534 should therefore be thought of as improving the domain V of LLΔ472 by preventing a dysfunctional two base loop forming rather than improving the stability of the structure. This interpretation also re-forms the 3 base pair stem (a). As a three base stem is probably formed in the very ts parent LLΔ472 virus it would appear that the length of stem (a) is less important to the functional domain V structure than the adjacent three base loop. The length of stem (b) would also appear to be less important as the change at 534 disrupted the 474/534 pair, leaving only six bases in the stem.

5.4.2 REVERTANTS OF LLΔ483

Plaques were also selected from the LLΔ483 virus when grown in L20B cells. Deletion of 483 shortens stem (b) from seven to six base pairs, leaving the U at 528 unpaired (see figure 5.6). The potential weak base pairing of the G at 484 with the U at 528 or the U at 514 would remove the 'hinge' base and reduce flexibility. A total of four plaques were picked at 37°C and one at 38°C. On assaying in L20B cells all the viruses were found to be less ts than the parent virus but more ts than Leon/Lansing (see table 5.7).

VIRUS	Log_{10} (pfu at 35°C/pfu at 37°C) in L20B cells	Log_{10} (pfu at 35°C/pfu at 38°C) in L20B cells
Leon/Lansing	0	0
LLΔ483	2.8	3.7
LLΔ483/37.1	0.4	1.1
LLΔ483/37.2	0.1	0.9
LLΔ483/37.3	0.2	1.1
LLΔ483/37.4	0.6	1.3
LLΔ483/38.1	0.5	1.1

TABLE 5.7 A comparison of log drop in titre at 37°C and 38°C as compared to that at 35°C for Leon/Lansing, LLΔ483 and the variants plaque picked from it, four were picked at 37°C and one was picked at 38°C.

Sequence data obtained from PCR products of all of these viruses showed that two mutations had arisen: a change of G to A at 482 and a change of G to A at 531. The domain V structure of the parent LLΔ483 virus and the theoretical structure of the variant are displayed in fig 5.6. The change at 531 can be interpreted as strengthening the 480/531 base pair from UG to UA which would strengthen stem (b) previously shortened by the deletion at 483. The change at 482 would destroy the base pair with 529, shorten the stem even more and enlarge the loop, originally made by bases 514-527, on both strands.

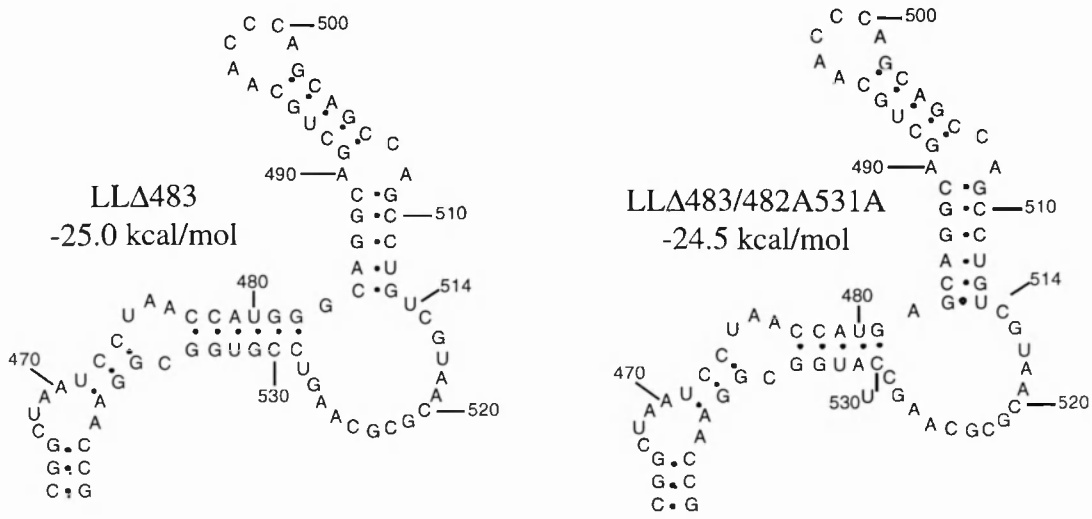


FIGURE 5.6 Probable structures of domain V in the virus LLΔ483 with a free energy of -25.0 kcal/mol and a variant of this with changes of G to A at 482 and a G to A at 531 and a free energy value of 24.5 kcal/mol (*c.f.* -27.5 kcal/mol for Leon domain V). Previous type 3 numbering is retained. Both structures were analysed using the 'foldRNA' and 'squiggles' programmes of GCG.

This suggests that the flexibility that the 'hinge' base and the loop between stems (b) and (c) allow is functionally more important than definite stem lengths. Placing an extra A at base 482 into the 'hinge' would allow the original G at 483 to pair with the U at 514, increasing the length of stem (c). In addition the original 1 base 'hinge' would re-form (see figure 5.6) and the potential for the G at 484 to base pair with the U at 528 is removed. The loop is already extended from the deletion at 483 and the slight rearrangement could alter the three dimensional folding or stacking of the bases in the loop.

The variants were still more ts than a Leon/Lansing virus, probably due to the shorter stem (b) and it would appear that this is a compromise. Obviously an insertion of a base at 483 would be the ultimate reversion to re-form the original domain V shape but this is very rarely seen as the result of replication errors, if at all. The structure formed therefore is not as efficient in its role as the Leon structure. Both the variant and parent domain V structures have higher free energy values than the Leon domain V at -27.5 kcal/mol, making them slightly more unstable thermodynamically. The variant however has a higher

free energy value than the parent by 0.5 kcal/mol indicating that the shape and formation of domain V is more important to the functional structure than the free energy. It should also be noted that the discovery of two mutations from one round of plating and selection of plaques is unusual but not totally impossible, particularly as each mutation would individually be advantageous.

5.4.3 REVERTANTS OF LLΔ483/528

The third virus containing a deletion in domain V had the base pair 483/528 removed and the resulting probable structure is displayed in fig 5.7. As can be seen, removal of this base pair resulted in a shortened stem (b) but the loop remained unchanged. Unlike LLΔ483 the removal of both 483 and 528 creates no new potential base pairing for the 'hinge' base at 484. Free energy of this is calculated as -24.2 kcal/mol (*c.f.* -27.5 kcal/mol for Leon).

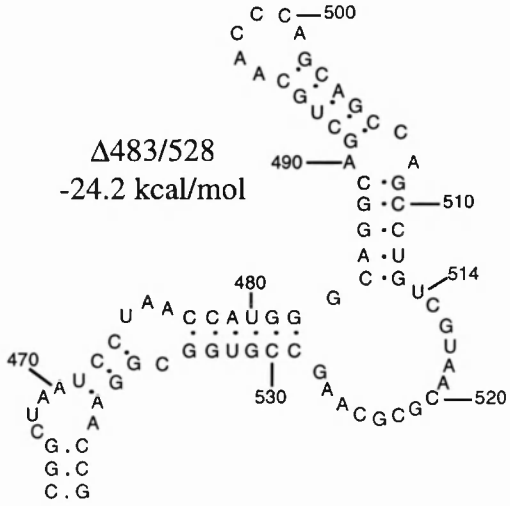


FIGURE 5.7 Probable structure of domain V in the virus LLΔ483/528. Previous type 3 numbering is retained. This structure was analysed using the 'foldRNA' and 'squiggles' programmes of GCG.

As above, LLΔ483/528 was grown in L20B cells under agar and plaques were picked after three days, four at 37°C and two at 38°C. Viruses grown from these plaques were assayed and all had similar phenotypes to the parent LLΔ483/528 virus (see table 5.8). In a second round of plaque selection from LLΔ483/528/37.1, three were picked from 37.5°C and two

from 38°C and on assaying they were found to be as ts as the parent virus (see table 5.8). This correlated with sequencing information which showed that viruses from the plaques contained no changes in domain V.

A

VIRUS PLAQUES PICKED IN FIRST ROUND	Log_{10} (pfu at 35°C/pfu at 38°C) in L20B cells	Log_{10} (pfu at 35°C/pfu at 38.5°C) in L20B cells
Leon/Lansing	0	0
LLΔ483/528	1.1	5.3
LLΔ483/528/37.1	1.2	3.2
LLΔ483/528/37.2	1.3	>5.3*
LLΔ483/528/37.3	1.2	>6*
LLΔ483/528/37.4	1	>5*
LLΔ483/528/38.1	1.4	>5.3*
LLΔ483/528/38.2	2	>5.6*

B

VIRUS PLAQUES PICKED IN SECOND ROUND	Log_{10} (pfu at 35°C/pfu at 37.5°C) in L20B cells	Log_{10} (pfu at 35°C/pfu at 38.5°C) in L20B cells
Leon/Lansing	0	0
LLΔ483/528	1.6	4.4
LLΔ483/528/375.1	1	>5
LLΔ483/528/375.2	1	5
LLΔ483/528/375.3	1	4.1
LLΔ483/528/38.1	0.8	5
LLΔ483/528/38.2	0.5	5

TABLE 5.8 A comparison of log drop in titre at various temperatures as compared to that at 35°C for Leon/Lansing, LLΔ483/528 and the plaques picked from it. Table A lists the data for the first round of plaque selection and table B lists the data for the second round of plaque selection.
* A finite number could not be measured from the dilutions used as the titre of the viruses used was too low to produce plaques at the higher temperatures.

These results suggest that although the LLΔ483/528 virus is ts and probably attenuated, domain V is genetically stable such that any improvement would entail more than one mutation and unstable or non functioning intermediate structures with only one mutation

are unlikely to form. This virus is also less ts than the LLΔ483 virus but as ts as the variants of LLΔ483 selected in L20B cells despite all three domain V structures having very similar values of free energy. This suggests that the extra base pairing potential of the 'hinge' base found in LLΔ483 but not LLΔ483/528 is detrimental to the function of domain V secondary structure. The selection pressure for any particular mutation in LLΔ483/528 could therefore be thought of as relatively small with any changes arising in the sequence giving the virus no advantage. This is in direct contrast to the LLΔ472 and LLΔ483 viruses from which variants arose after one round of plaque selection.

5.5 NON TS REVERTANTS OF A VIRUS WITH WEAKENED BASE PAIR MUTATIONS IN THE 5'NCR

In stem structures formed from double stranded RNA a CG base pair forms a slightly stronger bond than a AU base pair which is stronger than a UG base pair. In an aim to produce a more genetically stable vaccine strain it has been suggested that the 5'NCR domain V be engineered so as to weaken the structure in a way that the virus cannot easily revert whilst still rendering the virus attenuated (Macadam *et al.*, 1992). Therefore viruses with GC base pairs replaced by UA base pairs were made by mutagenesis to test this (personal communication, J. W. Almond & A. J. Macadam). The only way these viruses could revert is by mutating on both sides of the stem simultaneously which is unlikely as each single mutation would make a weaker or disrupted base pair.

5.5.1 LL472/537UA/478/533UA/480/531UA

A number of combinations to genetically stabilise domain V have been proposed for a possible future vaccine strain. One virus was used for selection of non ts revertants in L20B cells as described above. The virus investigated here was the result of site directed mutagenesis of a Leon/Lansing construction, LL472/537UA/478/533UA/480/531UA

(LL/UA/UA/UA) and had changes to three base pairs. The probable structure of this virus is displayed in fig 5.8 and this shows that one CG base pair in stem (a) and two base pairs in stem (b), one CG and UG, were mutated to UA.

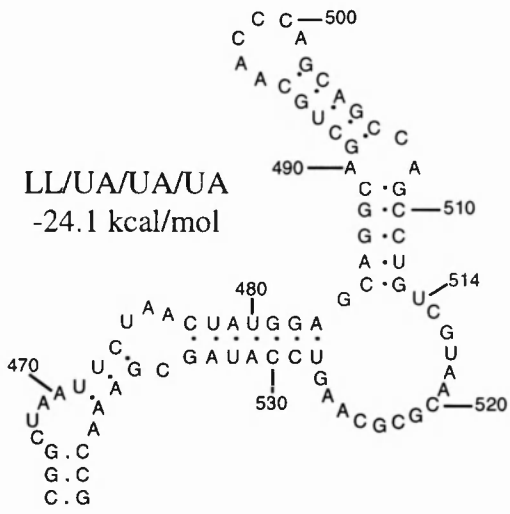


FIGURE 5.8 Probable structure of domain V in the virus LL472/537UA/478/533UA/480/531UA. The free energy value for the domain V of this virus was -24.1 kcal/mol (*c.f.* -27.5 kcal/mol for Leon). The structure was analysed using 'foldRNA' and 'squiggles' programmes of GCG.

Virus was grown in six well plates and initially four plaques were selected at 38.5°C. Results in table 5.9 show that the parent virus is very ts with a large drop in titre in L20B cells at 38.8°C compared to that at 35°C. Therefore changing base pairs from CG to UA appears to destabilise the domain V secondary structure to quite an extent. In addition, all of the viruses derived from the selected plaques showed no change in the ts phenotype and from sequencing data there were no mutations in domain V. Two of these plaque selected viruses, LL/UA/UA/UA/385.1 and LL/UA/UA/UA/385.2, underwent further plaque selection at 38.5°C and again information from both sequencing and assay in L20B cells showed that no changes had occurred (see table 5.9).

A

VIRUS PLAQUES PICKED IN FIRST ROUND	Log_{10}(pfu at 35°C/pfu at 38.3°C) in L20B cells	Log10(pfu at 35°C/pfu at 38.8°C) in BGM cells
Leon/Lansing	0.7	1.1
LL/UA/UA/UA	1.5	4.2
LL/UA/UA/UA/385.1	1.7	4.6
LL/UA/UA/UA/385.2	1.7	5
LL/UA/UA/UA/385.3	1.7	4.7
LL/UA/UA/UA/385.4	1.4	4

B

VIRUS PLAQUES PICKED IN SECOND ROUND	Log_{10}(pfu at 35°C/pfu at 37.2°C) in L20B cells	Log10(pfu at 35°C/pfu at 37.8°C) in BGM cells
Leon/Lansing	0	0
LL/UA/UA/UA	0.4	0.7
LL/UA/UA/UA/385.1.1	0.6	0.9
LL/UA/UA/UA/385.1.1	0.3	0.7

TABLE 5.9 A comparison of log drop in titre at various temperatures as compared to that at 35°C for Leon/Lansing, LL/UA/UA/UA and the plaques picked from it. Table A lists the data for the first round of plaque selection and table B lists the data for the second round of plaque selection.

This evidence indicates that the LL/UA/UA/UA virus is unable to revert easily in L20B cells. Much more work would need to be carried out with LL/UA/UA/UA before it was deemed unable to revert completely. The virus studied here is also of a Leon/Lansing construction. Any suitable changes would need to be introduced into Sabin vaccine strains and results from preliminary studies with a type 3 version are optimistic (personal communication, A. J. Macadam).

5.6 DISCUSSION

The results presented in this chapter follow on from results of chapter three. Previously the changes in 2A were found to be able to compensate for mutations in domain V that

disrupted secondary structure and render a virus non ts in BGM cells. However in a monkey neurovirulence test such non ts viruses remained attenuated. This suggested that the phenomenon was cell specific and involved interactions with cell specific factors, an observation that was substantiated when viruses were assayed in L20B cells as the changes in 2A had no compensating effect in these cells. Viruses were all found to be as ts as the parental viruses with only domain V mutations. Consistent with this all non ts revertant viruses selected in L20B cells were found to have mutated to re-form domain V secondary structure. None were found to have changes in 2A.

All the non ts revertants of Sabin 2 were found to have directly back mutated at 481. This meant that L20B cells did not exert any selection pressure on polioviruses to mutate in 2A and that these cells did not express the cellular factor necessary for this activity of 2A. Following on from this observation, it was decided that L20B cells could therefore be used as a tool in the study of domain V structure. Although it is possible to select non ts revertants in BGM cells which have reverted in domain V, the majority were found to have mutated in 2A instead. In addition, viruses with no obvious method of domain V reversion, *e.g.* a virus with a deletion at 472, could now be investigated. Such viruses were all constructed in the laboratory and were not naturally occurring but their reversion could give insight about important secondary or tertiary structures.

Viruses selected from LLΔ472 and LLΔ483 were less ts variants of the parent rather that non ts revertants. This illustrates to what extent base deletions disable the secondary structure of a virus. Two of the viruses were not found to alter on growth at elevated temperatures. Any changes that the virus is able to achieve would probably be even more detrimental to the virus. Therefore LLΔ483/528 and LL/UA/UA/UA appeared to be genetically stable in L20B cells. If LL/UA/UA/UA was confirmed to be attenuated then

either of these sequences could be incorporated into useful vaccine candidates unless such viruses were found to be over-attenuated.

Variants of LLΔ472 were found to have a mutation that favoured the formation of a three base loop between stems (a) and (b) at bases 474-476. Since this loop in the more stable parental structure was reduced to two bases in the parent virus the length of loop must therefore be an important factor in the secondary structure of domain V, possibly allowing correct folding of the overall structure. The re-formed loop in the variant also has a different sequence to that found in all polioviruses. This, along with the shorter stem (b), could contribute to the reason why the virus remains ts but it would appear that sequence is not as important as length. Both the length and sequence requirements of the loop could be further studied by the construction of site directed mutants.

Variants of LLΔ483 illustrated the importance of unpairing the 'hinge' base at position 484 to the functional structure of domain V. In the parent, the G at 484 could potentially base pair with the U at 514 or the U at 528, albeit weakly. The mutations found in the variant viruses were rationalised as forcing an unpaired A into the 'hinge' position whilst shortening stem (b) by one base pair. This allowed the original 'hinge' base to form a potential GU pair and indicated that the flexibility at the 'hinge' was more important than the specific length of stem (b). The slight rearrangement of the fourteen base loop and the stems either side of it could partly explain why the variant is still ts.

These results are consistent with the reversion of the Sabin 2 attenuating mutation at 481 where potential base pairing of the 'hinge' base, 481 A to 511 U is always weakened by mutation of 481 A to G. In all other polioviruses the 'hinge' base is a G already and has the potential to form a base pair with the U at 514, *e.g.* in Leon, but it is a relatively weak

pair. The requirements of the unpaired 'hinge' base could be investigated further by the construction of site directed mutants with different base pairing at this region.

In summary the results presented indicate a number of features about the compensatory role of 2A. Cell specificity means that direct interaction of 2A with the 5'NCR is ruled out. A cellular factor is instead required for this role. The large number of changes in 2A that are able to bring about the effect suggest that they disrupt a binding function, releasing the factor to itself interact with the 5'NCR. The factor would appear to absent or sufficiently different in L20B cells and the CNS such that the interaction with 2A is not disrupted or absent. This would mean the factor acts to enhance translation rather than as a canonical factor.

Further investigations into cell specificity of the changes in 2A could prove useful. Brief attempts at using the human diploid MRC-5 cell line, derived from foetal lung tissue (Jacobs *et al.*, 1970), showed that assay conditions needed to be refined and no results were obtained. The 2A mutations may have no effect in the cell lines used for vaccine manufacture as no significant mutations have been found in the past. The results presented in this chapter however do present L20B cells as a valuable tool in future experiments to identify significant cellular factors.

CHAPTER SIX

MUTATIONAL ANALYSIS OF BASES 474-476 IN DOMAIN V OF THE 5'NCR

6.1 INTRODUCTION

Selection of revertants of ts Leon/Lansing derived viruses in L20B cells was described in chapter five. The aim of this was to study secondary and tertiary structure within the 5'NCR of polioviruses. The viruses carried mutations in domain V which conferred a ts phenotype as well as affecting virulence of the virus. Previously the BGM cell line was used to select non ts revertants but due to selection pressures enforced by these cells most non ts revertant viruses were found to have compensating amino acid changes in the protease 2A. Although this highlighted a previously unknown role of this protein it provided no information on requirements of the 5'NCR.

Chapter five also outlined that the viruses with compensating mutations in 2A were still ts when assayed in L20B cells so that the ability of 2A to counteract the effect of domain V mutations appeared to be cell specific. This led to the use of L20B cells to select non ts revertants with 5'NCR mutations as it was presumed there would be no selection pressure on the virus to mutate in 2A.

The viruses used in this part of the study were unable to directly back mutate. One such virus was LLΔ472, a site directed mutant of Leon/Lansing with a deletion at 472 (Macadam *et al.*, 1992; Macadam *et al.*, 1994). This virus has the potential to form two slightly different structures as discussed in chapter five and these are shown in fig 6.1. The first structure is shown with the UAA loop between stems (a) and (b) and stem (a) reduced to two base pairs. The free energy of this is calculated as -23.2 kcal/mol (*c.f.* -27 kcal/mol for Leon). The second is shown with the loop reduced to AA whilst the U at 474 base pairs with the G at 536 allowing stem (a) to retain three base pairs. The free energy of this structure is -26.5 kcal/mol, making this the more likely structure to form.

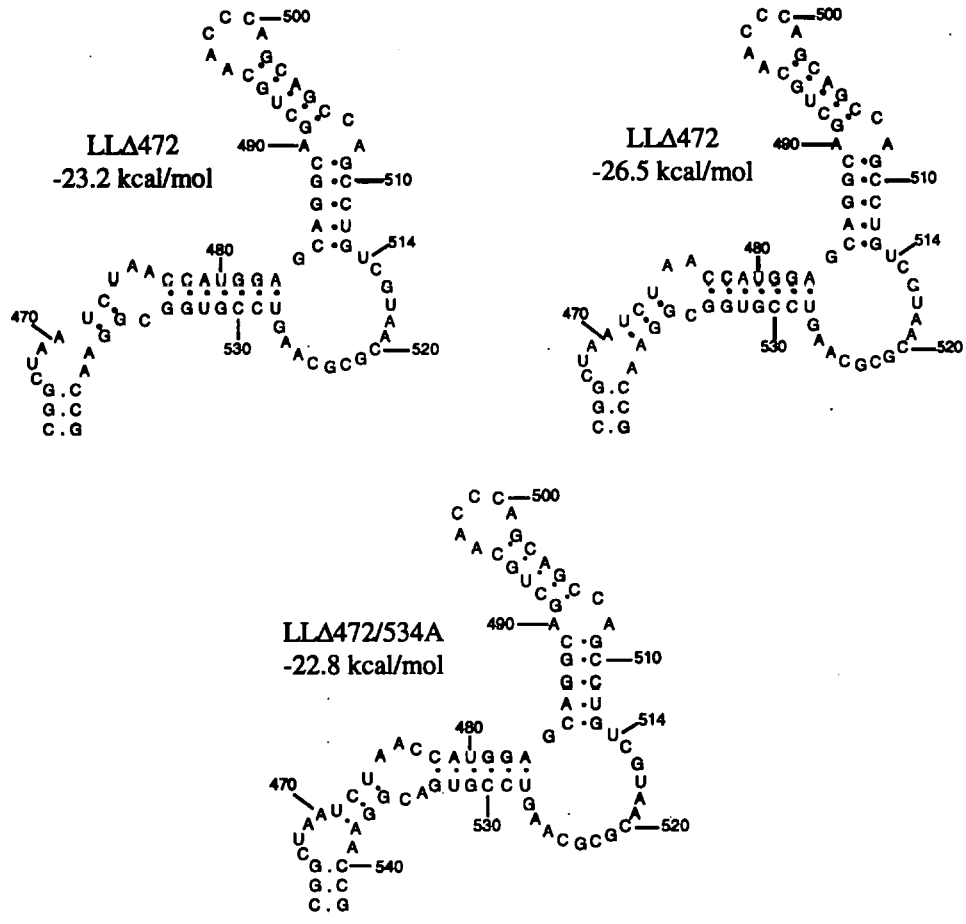


FIGURE 6.1 Two possible domain V structures of the parent LLΔ472 virus and the L20B derived LLΔ472/534A virus with the calculated free energy values. Structures and values were determined using 'foldRNA' and 'squiggles' GCG programs.

Less ts variants of LLΔ472 were found to have a change of G to A at 534. This was rationalised as reforming the three base loop between stems (a) and (b) and resulted in the reduction of stem (b) from seven base pairs to six (see fig 6.1). The sequence of the loop would become AAC. In addition stem (a) would be able to also reform as a three base stem. From this it could be concluded that the length or flexibility of this 474-476 loop is more important to the virus in terms of functional structure formed than the actual sequence of it or the length of stem (b). The change at 534 can be thought of as preventing the formation of a thermodynamically stable two base loop. The free energy value of the

variant, -22.8 kcal/mol, is higher than the free energy values for either form of the LLΔ472 domain V but the variant is less ts.

Data presented in fig 4.10 in chapter four showed that in all polioviruses the sequence 474-476 is conserved as UAA except in Mahoney and Sabin 1 where it is CAA. Furthermore the sequence here is conserved as NAA across the entire enterovirus group (Jackson *et al.*, 1994) although a wild type virus from Brazil was found to have a G at 476 (Minor & Dunn 1988). Therefore it would seem unlikely that many variations at these positions would result in viable viruses. Such high conservation may suggest that this loop interacts either with other areas of the 5'NCR, which may be discovered by mutant/revertant analysis, or with proteins involved in translation.

It was therefore decided that viruses with site-directed mutations in this 474-476 loop would be made to test the hypothesis. This chapter will describe how mutant viruses were constructed with changes in loop sequence and length. The effect of the mutations on the virus were analysed in terms of viability and growth at higher temperatures.

RESULTS

6.2 MUTATIONS TO BE INTRODUCED

It was decided to introduce three types of mutation into the loop formed by bases 474-476 in domain V of the 5'NCR in order to determine the requirements of length and sequence to the virus. The first type of mutation introduced any base into the three positions 474-476, the second deleted 476, and the third deleted bases 475-476.

6.2.1 STRATEGY FOR MUTAGENESIS

The mutagenesis strategy took advantage of the available unique restriction sites in the poliovirus sequence. Changes were introduced into DNA using mutagenic primers in a PCR reaction and amplification was kept to a minimum number of cycles to avoid the introduction of extra changes by the Taq polymerase. For this, it was necessary therefore to make a sub-clone in M13, inserting a portion of the 5'NCR from a clone of Leon/Lansing already in existence, pT7SFP (Skinner *et al.*, 1989). pT7SFP incorporated a T7 promoter such that RNA could be transcribed from the Leon/Lansing insert using T7 RNA polymerase.

Sub-cloning was achieved by cleavage of 1μg of pT7SFP with restriction enzymes KpnI (New England Biolabs) and SacI (New England Biolabs). SacI uniquely cleaves Leon/Lansing at position 751 while KpnI cleaves at 70 and 2267. After electrophoresis of the cleavage products a DNA fragment of 681 base pairs was purified from the agarose slice using the Prep-a-Gene method. The same procedure was carried out on 1μg of double stranded M13 (mp19 from New England Biolabs), uniquely cleaved at positions 6273 with KpnI and 6279 with SacI in the polylinker. The small fragment of six bases was lost by simple direct purification using the Prep-a-Gene method. The 681 base pair fragment of Leon DNA was then ligated into the large fragment of M13 using 0.5 units of ligase (Promega), incubating at room temperature.

This ligated M13 DNA was transfected into chemically competent TG1 cells and transformed bacteria were plated onto agar plates. White plaques formed from this were picked and phage grown in a TG1 culture at 37°C for extraction of DNA. Single stranded DNA was extracted for sequencing using the M13 1211 primer and then double stranded DNA was extracted for mutagenesis. The whole of the insert was sequenced to ensure that the sub-clone was correct.

After mutagenesis, described below, the M13 sub-clone and the original pT7SFP were then cleaved using MluI (New England Biolabs) and SacI (New England Biolabs), both of which uniquely cleave Leon/Lansing at positions 278 and 751 respectively. M13 has no MluI sites. The correct fragment from the M13 sub-clone, purified by using the Prep-a-Gene method on an excised gel slice, was then ligated into the large fragment purified in the same way from the cleaved pT7SFP again using 0.5units of ligase (Promega). Ligated pT7SFP derived DNA was then electroporated into competent DH5α cells and resulting bacteria spread onto selective agar. Colonies were picked and grown in culture at 37°C, DNA extracted and sequencing was carried out using primer αLL390 to check that the correct mutations were obtained. A schematic of the process described above is shown in fig 6.2.

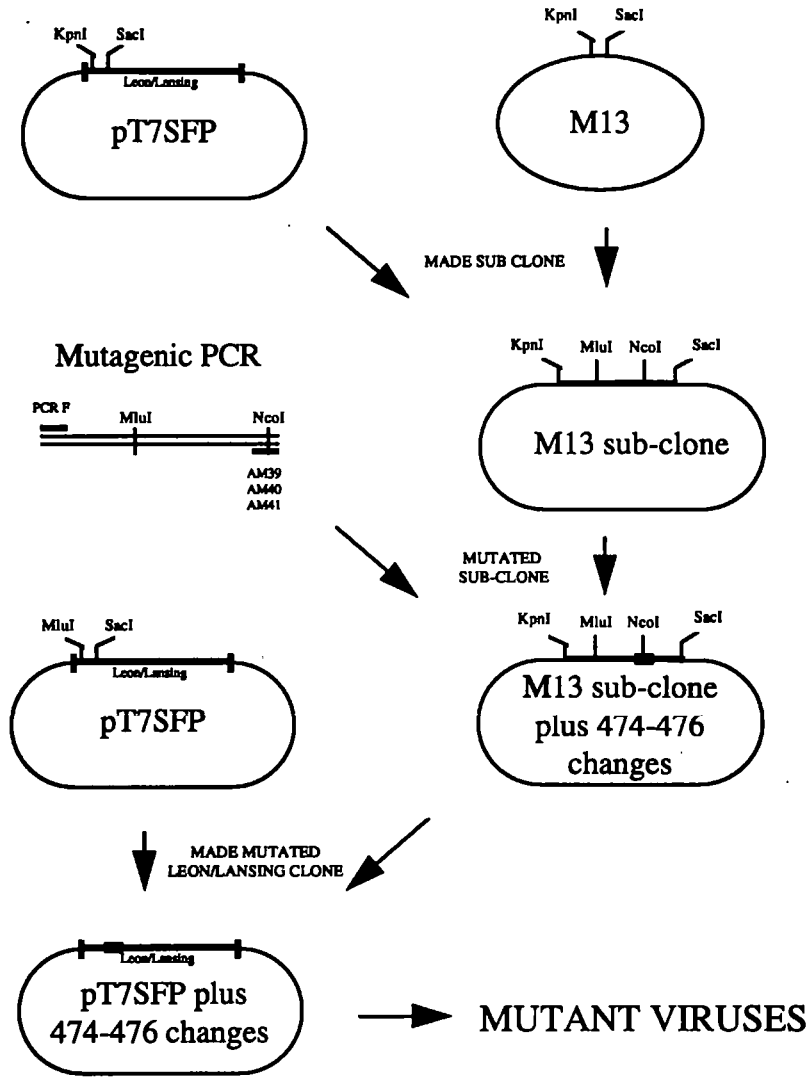


FIGURE 6.2 A schematic to show the strategy for constructing site-directed mutants of domain V at positions 474-476 of a Leon/Lansing virus. Initially a sub-clone was made in M13, mutated DNA from a PCR reaction was then ligated into the sub-clone and finally the mutated Leon/Lansing insert was ligated back into the Leon/Lansing full length clone.

RNA was made from correct clones using T7 RNA polymerase and plasmid DNA linearised with restriction enzyme SalI (New England Biolabs). This RNA was then transfected into HEp-2C and L20B monolayers, both in 25cm^2 flasks, to produce a liquid culture and in overlayed six well plates for selection of plaques.

6.2.2 MUTAGENESIS

Three mutagenic primers were synthesised (see chapter two). AM39 was made with degenerate bases at the positions corresponding to 474-476. AM40 was made with a deletion corresponding to base 476 and AM41 (see chapter two) was made with a double deletion corresponding to bases 475-476. All three were made complementary to the Leon sequence with at least twelve bases either side of the mismatches to enable accurate hybridisation to template DNA. In addition they incorporated the NcoI cleavage site at 477.

Using these antisense primers and PCR F (see chapter two) as sense primer, PCR reactions were carried out using a highly processive and accurate polymerase, Amplitaq (Perkin Elmer), on 2µg of Leon/Lansing cloned DNA. This enabled a minimum number of cycles to be used to limit PCR error. Products from these PCR reactions were purified using the Prep-a-Gene method prior to cleavage with restriction enzymes MluI and NcoI (New England Biolabs). These enzymes cleave Leon/Lansing at positions 272 and 477 respectively and cloning was performed as described above.

Using this approach, both the single and double deletions were successfully introduced into the M13 sub-clone. In addition, in a clone with a double deletion mutation, an extra change of C to U at position of 472 had been introduced presumably from a PCR error. It was decided that this mutation should be included in this experiment. In the case of sequence changes, from forty two plaques that were selected, eleven were found to have different sequence at equivalent Leon/Lansing positions 474-476. The M13 sub-clones with mutations in the Leon/Lansing insert that were obtained are summarised in table 6.1. Unfortunately not all of the M13 mutated sub-clones were reconstructed into the pT7SFP clone due to limited time factors.

M13 CLONE	MUTATIONS INTRODUCED	RECONSTRUCTED VIRUSES
M39.1	474-476 CGG	n.d.
M39.2	474-476 ACA	n.d.
M39.3	474-476 TTA	n.d.
M39.4	474-476 ACC	n.d.
M39.5	474-476 GGT	n.d.
M39.6	474-476 TTT	n.d.
M39.7	474-476 AAC	n.d.
M39.8	474-476 ACG	n.d.
M39.9	474-476 ATC	n.d.
M39.10	474-476 CAC	n.d.
M39.11	474-476 GAA	n.d.
M39.12	474–476 ATC	LL/AUC
M40.1	Δ476	LLΔ476
M41.1	ΔΔ475–476	LLΔΔ475–476
M41.2	ΔΔ475–476/473 T	LLΔΔ475–476/473 U

TABLE 6.1 A list of all the M13 sub-clones obtained after ligation of mutagenic PCR products.
n.d. means not done

6.3 EFFECTS OF Δ476 ON LEON/LANSING

The Leon/Lansing insert from sub-clone M40.1 was substituted into PT7SFP using restriction enzymes MluI and SacI as described above. The following transfection of RNA from this clone was carried out in triplicate but results were the same for each. The success of a transfection relies on the quality and quantity of the RNA produced. As RNA is unstable, the failure to produce virus may not always be due to its non viability. Repeating this in triplicate therefore confirms that the results observed are valid.

Plasmid DNA was extracted from bacteria, checked by sequencing with primer αLL390 and linearised with SalI. Full length RNA was then transcribed from this using T7 RNA polymerase. One quarter of the RNA was used to transfect a HEp-2C monolayer in a $25cm^2$ flask and one quarter was used to transfect a L20B monolayer in a $25cm^2$ flask. These and control mock transfected flasks were then incubated at 35°C and checked daily for CPE. This would produce a stock of any virus which grew. The rest of the RNA was used to make serial log_{10} fold dilutions in the 1 X HBSS/glucose/DEAE-dextran mix for transfection of L20B and HEp-2C monolayers in a six well plate. These were overlayed and incubated at 35°C for three days before visualising plaques with neutral red for plaque selection.

All three duplicate experiments for this plasmid gave the same results. While LLΔ476 was viable in HEp-2C cells, no virus grew in L20B cells in flasks or under overlay. This was probably due to the fact that these cells are less permissive than HEp-2C cells to growth of viruses with detrimental mutations in domain V of the 5'NCR (Macadam *et al.*, 1992). The HEp-2C infected flasks showed CPE after three to four days whilst the control flask transfected with RNA from the pT7SFP clone showed CPE after only twenty four hours. The three or four days taken for signs of CPE is relatively slow and would either indicate poor quality RNA from the T7 RNA polymerase reaction or that the growth of the virus was severely retarded. Plaques were visualised in the HEp-2C infected six well plates only in the wells inoculated at a 2 log_{10} dilution of the original RNA whereas the plates transfected with RNA from pT7SFP had plaques in the wells inoculated at a 4 log dilution. This probably indicated again that the RNA was of poor quality or quantity or that the virus was retarded in growth. Three plaques were picked from the HEp-2C plates: LLΔ476/35.1; LLΔ476/35.2 and LLΔ476/35.3. Stocks were subsequently grown in HEp-2C monolayers in $25cm^2$ flasks. RNA extracted from these viruses was made into cDNA and amplified in a PCR reaction using primers PCR F and PCR 9. Sequencing using

primer αLL390 (see chapter two) in a T7 reaction confirmed that these viruses had the deletion at 476 (see fig 6.3).

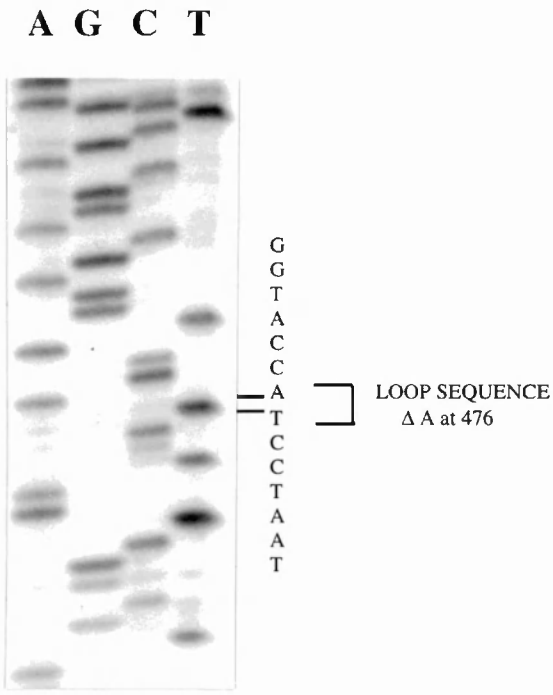


FIGURE 6.3 Photograph of sequencing radiograph showing the base deletion at position 476 for LLΔ476. Sequencing was carried out on a PCR product made from primers PCR F and PCR 9 and primer αLL390 was used in a T7 sequencing reaction. Lanes are marked AGCT and the position of the deletion is indicated.

Viruses derived from the selected plaques were assayed in both BGM and L20B cells. The results from this are listed in table 6.2. Viruses appear to be viable although all display a severe ts phenotype in both BGM and L20B cells compared to Leon/Lansing. The ts phenotype of the LLΔ476 viruses was more severe in L20B cells than in BGM cells. This confirms previous observations that viruses will grow at higher temperatures in BGM cells than in L20B cells (Macadam *et al.*, 1992). In L20B cells, plaques were relatively small at 35°C and titres were reduced by at least three logs at 37.5°C. It is less surprising therefore that transfection of the RNA into L20B cells produced no virus. The HEp-2C flask derived

stocks of the original transfections were also assayed in BGM cells. However these stocks had such low titres at 35°C that it was impossible to measure the log drop in titre at 38°C and 39°C. These viruses and the plaque derived viruses should have been assayed at lower temperatures so that a direct comparison could be made if time had permitted.

VIRUS	Log_{10}(pfu at 35°C/pfu at T°C) for BGM cells		Log_{10}(pfu at 35°C/pfu at T°C) for L20B cells	
	38°C	39°C	37.5°C	38.5°C
Leon/Lansing	0.4	1.7	0.2	0.4
LLΔ476/35.1	2.8	4.0	3.0	4.9
LLΔ476/35.2	2.2	3.5	3.2	3.7
LLΔ476/35.3	2.7	4.2	3.1	>4.0

TABLE 6.2 List of titre drop with temperature of LLΔ476 viruses derived from plaques measured as the ratio of pfu at 35°C to that at 38°C and 39°C for the BGM assay and at 37°C and 38°C for the L20B assay.

As for previous chapters the structures and free energies discussed here were calculated using the GCG (GCG Program Manual 1994) programs 'foldRNA' and 'squiggles'. The probable structure of domain V of the LLΔ476 virus is displayed in fig 6.4. As before, the loop formed by bases 514 to 527 and base 484 (type 3 numbering) was forced to be unpaired following the structure of Skinner *et al.*, (1989) and Pilipenko *et al.*, (1989).

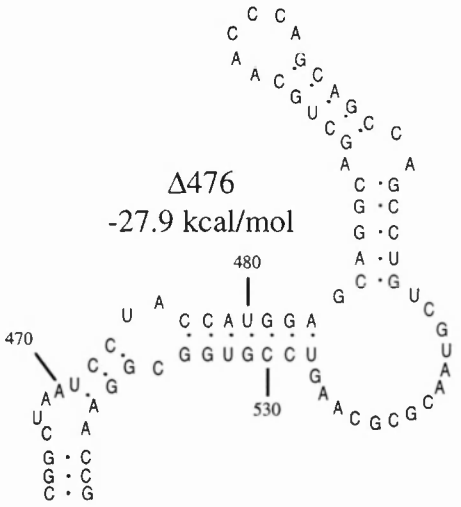


FIGURE 6.4 Probable structure of the domain V of virus LLΔ476 with the free energy values of -27.9 kcal/mol (*c.f.* -27.5 kcal/mol for Leon). The structures and energies were calculated using GCG programs 'foldRNA' and 'squiggles'. Numbering is of the original type 3.

Figure 6.4 shows that the deletion at 476 shortened the loop from 3 to 2 bases. The virus could restore this by unpairing bases 477/534 and shortening the stem (b) which was thought to have occurred in the L20B derived revertant of LLΔ472 in chapter five. In the same way, unpairing base pair 473/536 would also restore the loop by shortening stem (a). However these changes would unpair a strong GC pair and either unpairing would increase the free energy by 1.7 kcal/mol. Alternatively, the virus could unpair bases 473/536 and force the C at 473 to become part of the loop. This would cause a rearrangement and shortening of the stem (a), pairing 471 with 537 and 472 with 536. This second structure however has a higher free energy than the first by 3.3 kcal/mol. Therefore the virus probably forms the structure with the shorter loop, having the lowest free energy value and it is this disruption to the secondary structure which presumably causes the ts phenotype. Further analysis of structure and function can be addressed from changes that might occur from selection of non ts revertants.

6.4 LLΔΔ475–476 AND LLΔΔ475–476/473 U

The inserts of M13 sub-clones M41.1 and M41.2 were cloned into PT7SFP in the same way as the insert of M40.1 described above. The resulting plasmids were sequenced before the production of full length RNA using T7 RNA polymerase (see fig 6.5). Transfection was carried out in the same way as for LLΔ476, in flasks and six well plates of HEp-2C cells and L20B cells, except that for these, transcription and transfection were done in duplicate.

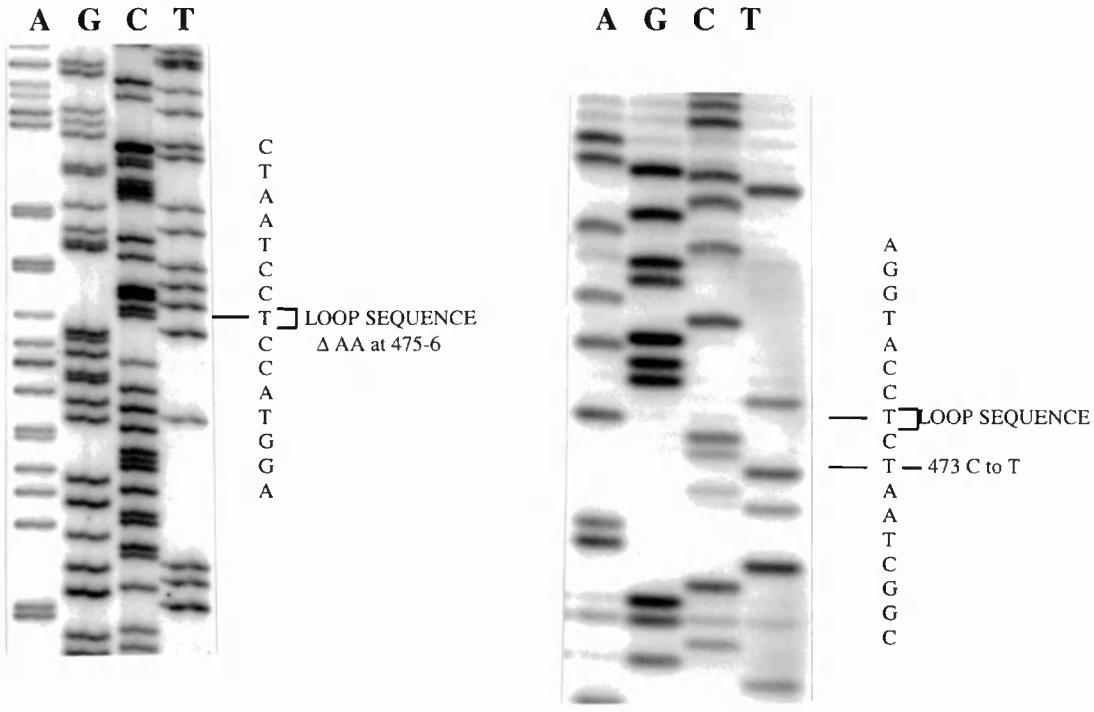


FIGURE 6.5 Photographs of sequencing autoradiographs showing the base deletions at 475 and 476 for LLΔ475–476 on the left and the additional change at 473 for LLΔΔ475–476/473 U on the right. T7 sequencing was carried out on plasmid DNA. Primer 13/II was used LLΔ475–476 and αLL390 was used for LLΔΔ475–476/473 U. Lanes are marked AGCT and the position of the deletions and changed are indicated.

For both LLΔ475–476 and LLΔΔ475–476/473 U no plaques were formed in plates and no CPE was observed in either HEp-2C or L20B cells in flasks. Positive control flasks showed CPE after 24-48 hours. To test for the presence of virus flasks showing no CPE were frozen after seven days and tissue culture fluid from these flasks was plaqued in both BGM

and L20B cells. No plaques were observed at 35°C. In addition, RNA extraction followed by cDNA synthesis and PCR amplification was carried out and no PCR product was obtained. It appeared therefore that these mutations in domain V resulted in non viable viruses.

The probable structure of LLΔΔ475–476 is displayed in fig 6.6. The loop is clearly reduced and would inevitably disrupt the overall folding of domain V and the rest of the 5'NCR. If the base at 473 is forced to become part of the loop and the stem forced to pair as 471/537, the energy of the structure is increased by 4.8 kcal/mol. This severe disruption to domain V could explain why the virus was non viable although the structure has a relatively low free energy value.

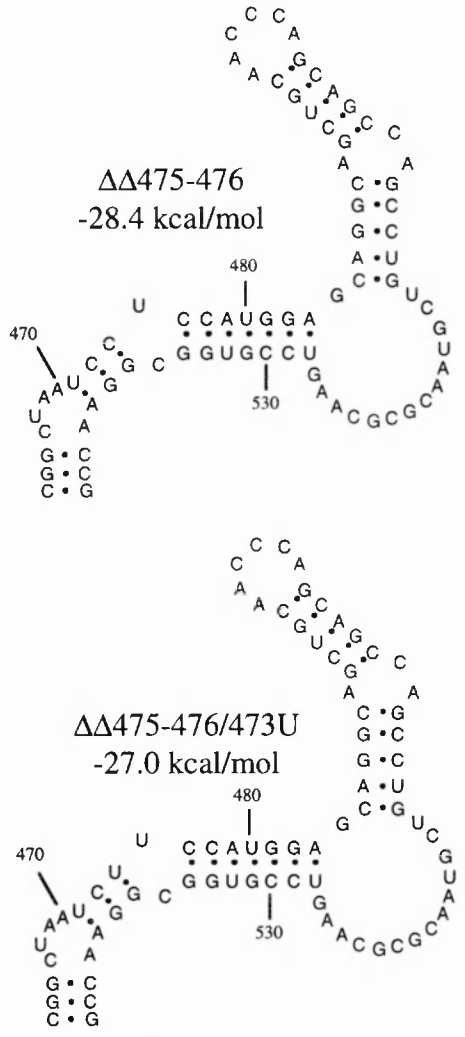


FIGURE 6.6 Probable structures of 5'NCR domain V regions of viruses LLΔΔ475–476 and LLΔΔ475-476/473 U with free energy values (*c.f.* -27.5 kcal/mol for Leon). Structures and energies were calculated using GCG programs 'foldRNA' and 'squiggles'. Numbering is of the original type 3.

The probable structure of LLΔΔ475–476/473 U is also displayed in fig 6.6. Again the structure is severely disrupted due to the reduction in loop length, despite the U at 473. In addition, forcing this base into the loop and rearranging stem (a) increases the energy of the structure by 3.4 kcal/mol. This may explain why the virus was non viable.

6.5 MUTATIONS OF 474-476 LOOP SEQUENCE

6.5.1 LL/AUC

The insert of M39.12 was cloned into PT7SFP as described above. Transfection from RNA made from the plasmid using T7 RNA polymerase was carried out as above,

transfecting into HEp-2C cells and L20B cells, in both flasks and six well plates. This was also only performed once.

This virus was found to be viable. CPE was observed in flasks after twenty four hours in both cell lines and plaques were found to form in the wells inoculated with RNA diluted to 4 logs of the original concentration. RNA was extracted from the supernatant of the HEp-2C flask stock for synthesis of cDNA. PCR amplification of this was carried out using primers PCR F and PCR 9 and purified fragments were sequenced using αLL390 as primer in a T7 reaction to confirm the correct sequence (see fig 6.7).

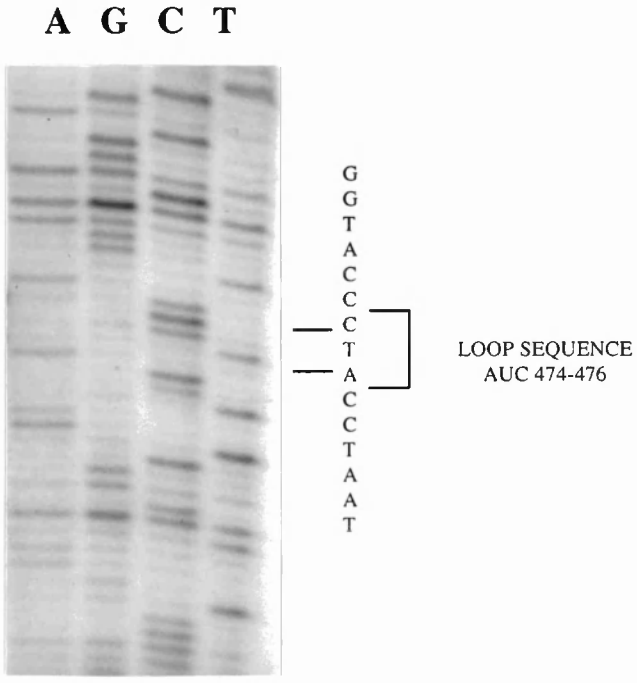


FIGURE 6.7 Photograph of the sequencing autoradiograph showing the base substitutions at 474-476 for LL/AUC. Sequencing was carried out on a PCR product made from primers PCR F and PCR 9 and primer αLL390 was used in T7 a sequencing reaction. Lanes are marked AGCT and the position of the base changes are indicated.

Virus from the HEp-2C stock was assayed in both BGM and L20B cells and the results can be seen below in table 6.3. Data shows that although LL/AUC is viable it displays a slight ts phenotype in both cell lines. As expected, the ts phenotype is more pronounced in L20B

cells in comparison with Leon/Lansing for each assay. But in comparison with LLΔ476 viruses in table 6.2, the LL/AUC virus is much less ts.

VIRUS	Log_{10}(pfu at 35°C/pfu at T°C) for BGM cells		Log_{10}(pfu at 35°C/pfu at T°C) for L20B cells	
	38°C	39°C	37.5°C	38.5°C
Leon/Lansing	0.4	1.7	0.2	0.4
LL/AUC	0.7	2.4	0.2	1.1

TABLE 6.3 List of titre drop with temperature of LL/AUC measured as the ratio of pfu at 35°C to that at 38°C and 39°C for the BGM assay and at 37°C and 38°C for the L20B assay.

Mutations that changes the loop sequence from UAA to AUC are therefore less detrimental to the domain V than a mutation that shortens the loop structure in terms of ts. Selection of the LLΔ472 variant that had a loop sequence of AAC also suggested that the sequence was not of fundamental importance to viability of virus. In addition transfection of RNA into L20B cells produced virus. These cells are much less permissive to viruses with detrimental growth mutations so this also confirms that a change of loop sequence to AUC has little effect on the virus. The results presented here further support that the length of loop is the main consideration. As the free energy of domain V of this virus is the same as that for Leon, the fact that LL/AUC shows a slight ts phenotype could be due to stacking effects exerted by these bases and interaction with adjacent bases. Again further analysis of structure and function would be obtained from changes occurring from selection of non ts revertants. The probable structure of domain V of LL/AUC is shown in fig 6.8.

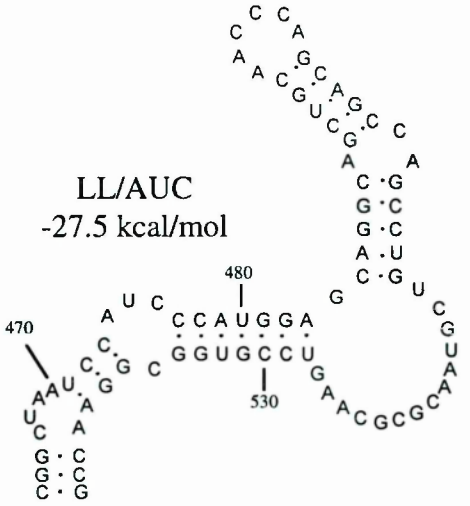


FIGURE 6.8 The probable structure of LL/AUC as calculated using 'foldRNA' and 'squiggles' programs of GCG. The free energy is calculated as -27.5 kcal/mol (*c.f.* -27.5 kcal/mol for Leon). Numbering is of the original type 3.

6.5.2 OTHER LOOP MUTANTS

Due to limitations on time none of the other inserts from the mutated M13 sub-clones were re-introduced into the Leon/Lansing clone. Consequently the full effects of changing the sequence of the loop 474 to 476 were not studied. Examples of sequencing gels showing mutations are displayed in fig 6.9.

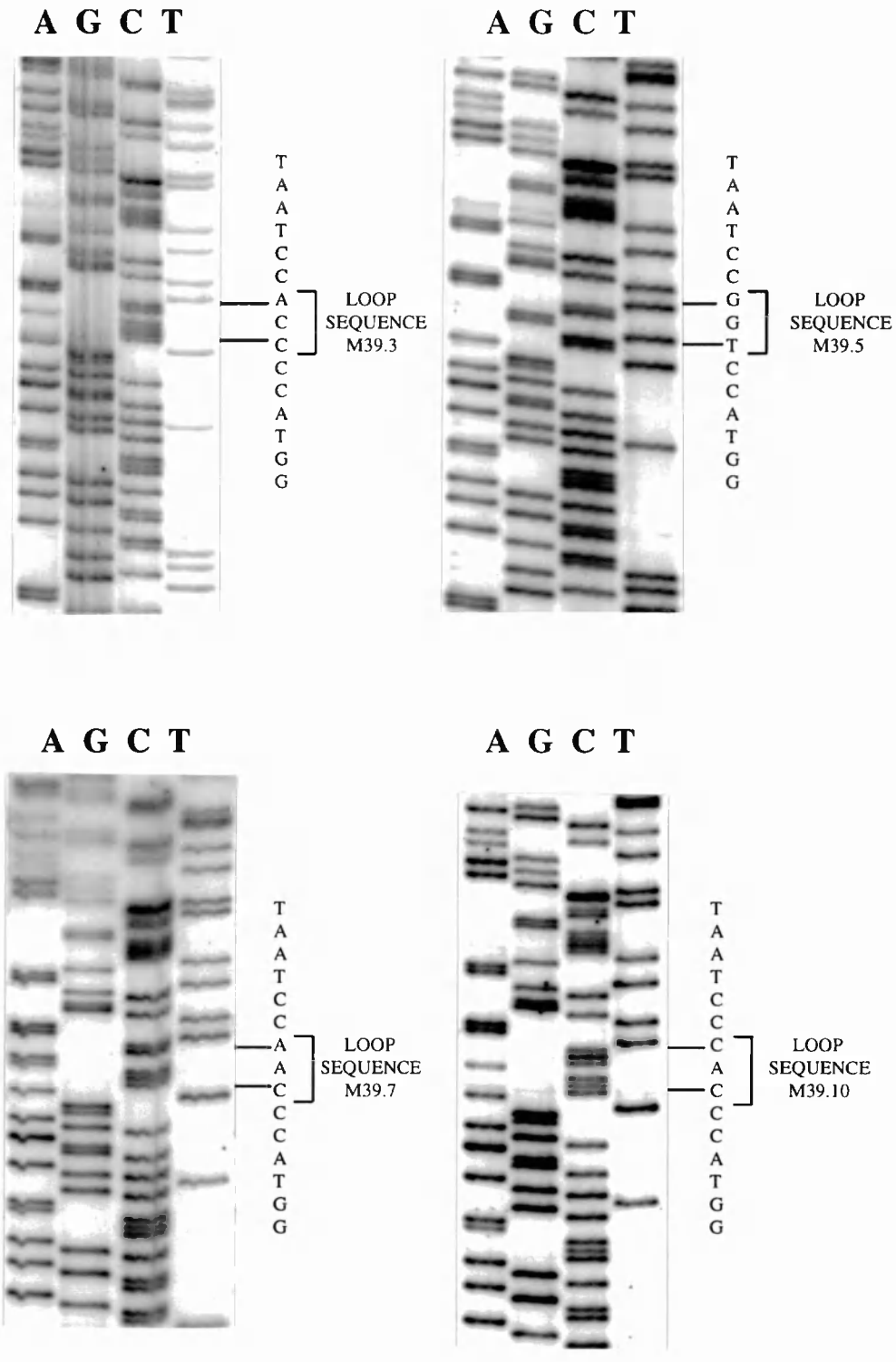


FIGURE 6.9 Photographs of sequencing autoradiographs of some of the M13 clones made with changes in the equivalent loop sequence at 474-476. Single stranded T7 sequencing was carried out using M13 1211 primer. Lanes are marked AGCT.

Tertiary RNA or protein interactions with this loop sequence are unlikely since LL/AUC and the variant of LLΔ472 were both viable. However some sequences of the loop may not be accommodated by the virus if they favoured alternative folding that changed the loop length. For example, any virus with a G in positions 474 or 476 would have the potential to reduce the loop to two bases. By pairing with the C at 535, the three base stem or the seven base stem would be lengthened. Such rearrangements may have implications on the whole folding of domain V and render a virus ts or non viable. Such considerations may explain the high conservation of the loop sequence. Alternatively the decrease in length of loop may be compensated by a decrease in a stem, possibly even in some kind of resonance of base pairing, *e.g.* a base pair of 474/535 could unpair 477/534. A wild type virus was found with a G at 475 (Minor & Dunn 1988) although as it was in the middle of the loop it is unlikely to have undergone the rearrangement described above.

Reconstruction of the inserts from M39.1, M39.5, M39.8 and M39.11 sub-clones all have at least one G in the equivalent 474-476 positions and these would be of particular interest. The predicted domain V structures of M39.1 and M39.5 are shown in figure 6.10. Both structures, M39.1 increasing stem (a) and M39.5 increasing stem 7(b), have free energy values of -29.4 kcal/mol (*c.f.* -27.5 kcal/mol for Leon) making them energetically very stable.

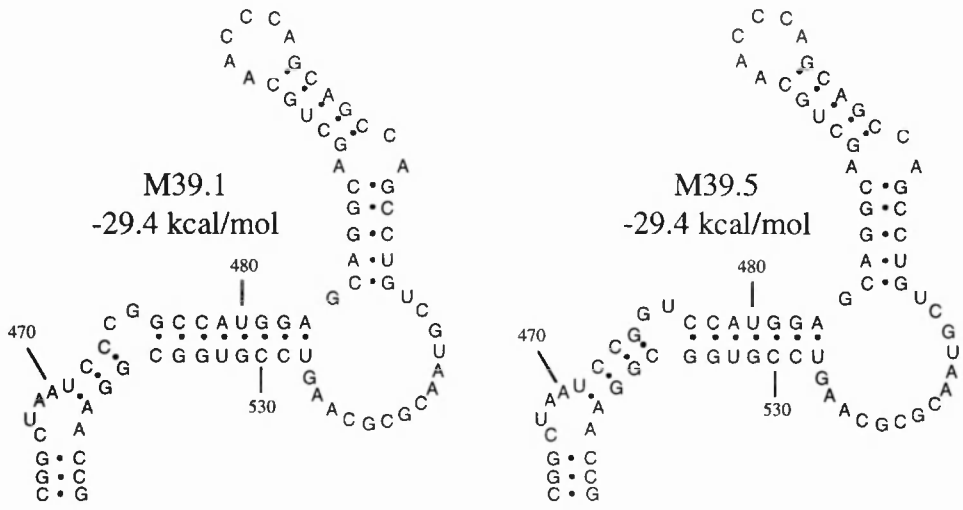


FIGURE 6.10 Predicted domain V structures that would result from the mutations introduced into the sub-clones M39.1 and M39.5. Free energy values are also shown (*c.f.* -27.5 kcal/mol for Leon). Structures and energies were calculated using 'foldRNA' and 'squiggles' GCG programs.

The insert of sub-clone M39.7 has the same sequence at 474-476 as that found in type 1 Mahoney and Sabin 1 viruses. It would be surprising if inserting this sequence into a Leon type domain V was detrimental to the virus. However there are some differences in domain V sequences between Leon and the type 1 viruses (fig 4.10 in chapter four). If the 474-476 loop did interact within the 5'NCR, a sequence of CAA at these positions in a Leon type background may disrupt the virus. Alternatively interaction may only involve the AA at 475-476 in which case any base could be tolerated at 474. Reconstruction of a virus from the insert of M39.7 would therefore be of interest.

6.6 DISCUSSION

Results from selection of variants from LLΔ472 in chapter five suggested that the length of the loop comprising bases 474 to 476 (type 3 numbering) was significant whilst the actual sequence was not. Loops would therefore appear to be as important to the functional structure of domain V as stems. The sequence of this loop is particularly highly conserved in enteroviruses which would suggest a vital role in the lifecycle of the virus rather than

simply acting as a spacer. Indeed many of the loop sequences in the secondary structure of the 5'NCR have highly conserved sequences.

The results presented in this chapter give further support to the importance of the length of the three base loop. Quite clearly the deletion of a single base is detrimental to the virus making it ts. From analysis of the domain V secondary structure the virus cannot easily re-form a three base stem from the disruption of bases in either stem (a) or (b) without significantly raising the free energy. In addition to this, the viruses made with two deletions in the loop were non viable. It is possible that the viruses may have been able to grow if incubated at a lower temperature. A three base loop in this position may therefore allow the correct folding of the domain to occur. A change in loop length could affect the overall shape of the domain to a greater extent than a change in stem length. Domain V is always depicted as being flat but in reality it is likely to have a three dimensional shape. This shape is thought to interact with other domains to form the IRES required for initiation of protein synthesis. There would be some limit as to length of stem but changes in lengths of loop could alter the folding and reduce the efficiency of initiation of protein synthesis.

The selection of revertants of LLΔ476 may provide more information of the sequence requirements of domain V. Non ts revertant viruses selected in BGM cells would be expected to have changes in 2A and previous experience is such that a deletion is very difficult to revert by base substitution. From the results above, if revertants were selected in L20B cells they would have a three base loop formed from the unpairing of the adjacent base pair in either stem (a) or (b).

The results presented here also suggest that changes in the sequence of the loop can be accommodated by the virus. A virus with an AUC loop is slightly more ts than a virus with a UAA loop which suggests that folding may have been affected by the introduction of steric hindrance. This was a little surprising as the high sequence conservation indicated a specific role for the sequence in tertiary structure or in protein interactions. The reconstruction of the other loop sequences, currently available as inserts into an M13 sub- clone, could confirm these observations. The behaviour of the sequence constructed in sub clones M39.1 and M39.5 would be especially interesting as the G bases allow rearrangements to occur which shorten the loop and probably result in ts viruses. In a further investigation of the requirements of loop length, revertants selected from these would be predicted to have mutations that allow the three base loop to be re-formed.

CHAPTER SEVEN

REQUIREMENTS OF BASE PAIRS 483/528 AND 484/514 TO DOMAIN V OF THE 5'NCR

7.1 INTRODUCTION

Results from previous chapters have highlighted the importance of unpaired bases in loops on the flexibility or folding of the domain V secondary structure and chapter six addressed the requirements of the loop made up of bases 474-476. Other results have shown the significance of the large fourteen base loop made up of bases 514-527 and the 'hinge' in the opposite strand, formed by base 484 (type 3 numbering, see fig 7.1). A number of viruses with changes or deletions around this area have been discussed already and revertants selected from them have given some insight into the importance of secondary structure formed.

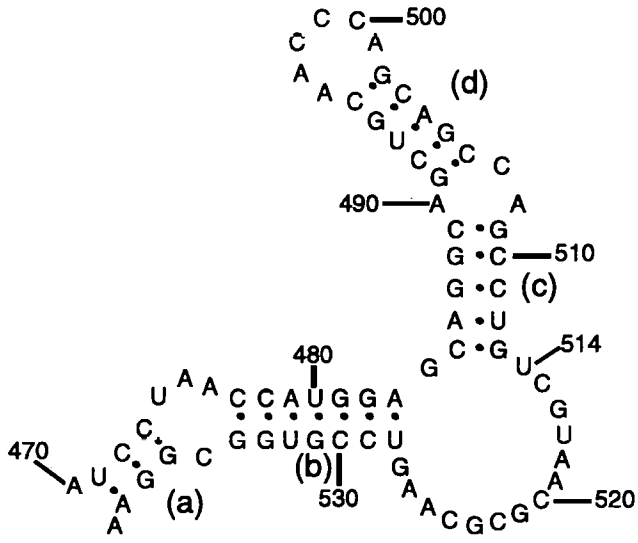


FIGURE 7.1 Predicted structure of domain V of the 5'NCR of polioviruses. Sequencing and numbering is of type 3 Leon virus. The four stems under discussion are named (a), (b), (c) and (d) for ease.

As discussed before, the rationale of the attenuating change found in domain V of the 5'NCR of Sabin 2 at 481 (484 in type 3 numbering) is not obvious. The base at position 484 is highly conserved as G (see chapter four, fig 4.10) and can potentially pair with the U at 514 (see fig 7.1). However this is not likely to occur as allowing 484 to pair with 514 in the Leon domain V secondary structure (whilst forcing the fourteen base loop between stems (b) and (c) to unpair) using 'foldRNA' on GCG raises the free energy by 5.4

kcal/mol. The G is therefore always shown as unpaired, able to form a 'hinge' at the loop between stems (b) and (c).

A change of G to A at 481 (484 in type 3 numbering) found in Sabin 2 is thought to favour the formation of a base pair with the U at 514 (Macadam *et al.*, 1991b; Macadam *et al.*, 1993). The 'hinge' nucleotide is therefore removed, the loop is reduced to thirteen bases and stem (d) increases from five to six base pairs. The ts and attenuated phenotypes are thought to arise from the 484/514 base pair altering the folding of domain V and raising the free energy of the functional structure. In revertant viruses a G at 484 is always selected in tissue culture and on passage through the gut, weakening the potential base pair with the U at 514. No other bases have been found at this position but this probably reflects the requirement for the base at 484 to be totally unpaired as a C or U would allow an alternative base pairing with the G at 527.

In chapter five the selection of a less ts variant of the virus LLΔ483 in L20B cells was described. LLΔ483 was originally produced by site directed mutagenesis of a Leon/Lansing clone and had a deletion at 483 (Macadam *et al.*, 1994). This deletion resulted in a domain V structure with a free energy value of -25.0 kcal/mol, a shortened stem (b) and an enlarged loop between stems (b) and (c) (see figure 7.2). As discussed in chapter five, the G at 484 can potentially make two GU base pairs, with either the U at 514 or the U at 528. If either of these base pairs form, the predicted free energy value for the structure is raised by 3.7 kcal/mol and the 'hinge' base is lost. The significance of these potential base pairs is illustrated by the less severe ts phenotype of LLΔ483/528 compared to LLΔ483 (see chapter five). The double deletant has a shortened stem (b) but only one potential base pair partner for 484, as in the wild type sequence.

The variant of LLΔ483 selected in L20B cells had two mutations, a change of G to A at 482 and a change of G to A at 531. In figure 7.2, the probable domain V structure of both LLΔ483 and the variant are shown. The change at 482 may appear surprising as it destroys the base pair 482/529 and reduces stem (b) even further to five base pairs. However, this is compensated to some extent by the change at 531 which strengthens stem (b), changing a GU pair to a stronger AU pair. In addition, the A at 482 can then form the 'hinge' between stems (b) and (c) whilst allowing the original G in the 'hinge' to base pair with the U at 514, increasing stem (c). This meant that the flexibility at the loop was the main consideration in this area of domain V rather than the precise lengths of stems (b) or (c).

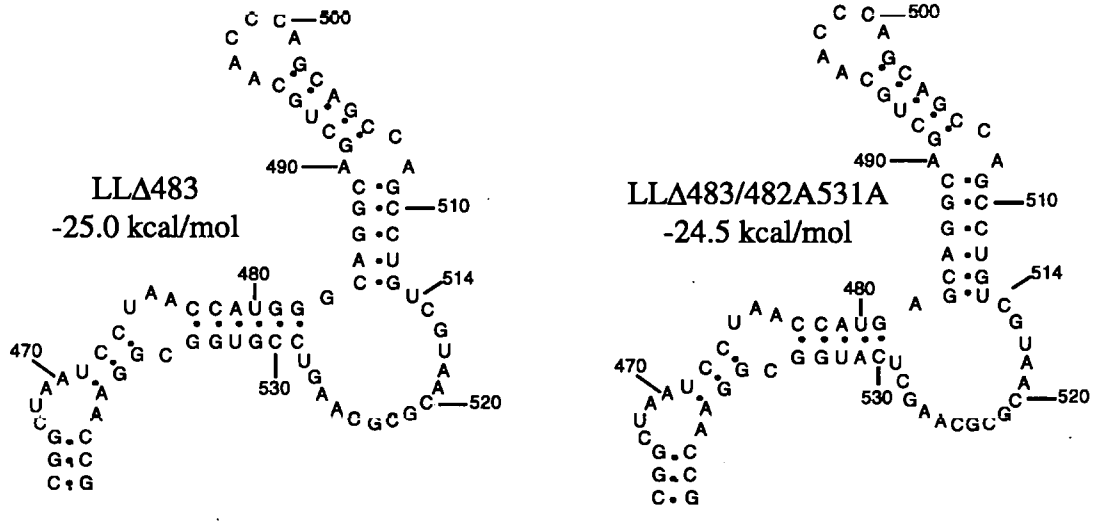


FIGURE 7.2 Probable structures of domain V of the 5'NCR in the virus LLΔ483 and a variant of this with changes of G to A at 482 and a G to A at 531. Both structures were calculated using 'foldRNA' and 'squiggles' programs of GCG.

In chapter three non ts revertants with changes in 2A selected in BGM cells from the virus LL514C were described. The parent virus, made by site directed mutagenesis of a Leon/Lansing clone (Skinner *et al.*, 1989), was shown to be ts in BGM cells at 39°C (see chapter three, fig 3.3) and a 2A amino acid substitution was found to counteract this. The change of U to C at 514 in domain V of the 5'NCR would allow a GC base pair at 484/514 to form. This could be seen as equivalent to the effect of a G to A change at 484 in Sabin 2

and result in a similar change of folding in this area. Similarly LL484/514 AU was found to be ts (personal communication, A. J. Macadam). The G to A change at 484, as in Sabin 2 allows a base pair of AU to form between 484 and 514.

Recently other clones of Leon/Lansing have been made by site directed mutagenesis of a Leon/Lansing clone by colleagues at Reading University. Changes at 483, 484, 514 and 528 were introduced into the sequence in a number of combinations. Viruses generated from them would have different abilities to form base pairs at 483/528 and 484/514 and therefore may give insight as to the importance of base pairing in this region. The clones produced are listed below in table 7.1 along with other viruses already partly characterised with differences at the same bases. In addition any revertants selected from ts viruses would give information on secondary structure of domain V.

A

CLONES	BASES AT 483/528	BASES AT 484/514
LL/GC/CC	GC	CC
LL/CA/CG	CA	CG
LL/CC/CC	CC	CC
LL/CG/CA	CG	CA
LL/CA/UC	CA	UC
LL/AG/GC	AG	GC
LL/CC/UC	CC	UC

B

VIRUS	BASES AT 483/528	BASES AT 484/514
LEON/LANSING, LL/AU/GU	AU	GU
LL514C, LL/AU/GC	AU	GC
LL484/514 AU, LL/AU/AU	AU	AU

TABLE 7.1 Table A lists the clones and potential viruses of Leon/Lansing with changes at 483, 484, 514 and 528. Table B lists the Leon/Lansing viruses already in existence with changes at these positions.

This chapter describes the generation of viruses with various changes at 483, 484, 514 and 528 from these clones. The initial viability of the virus and the ts phenotype as measured in a plaque assay was used to determine the importance of sequence and structure of domain V at this 'hinge' area. The phenotypes of LL/AU/GC and LL/AU/AU previously recorded will also be discussed in terms of potential base pairing at 483/514 and 484/528.

RESULTS

7.2 RECONSTRUCTION OF VIRUSES

Plasmid DNA for each of the eleven clones was linearised using SalI (New England Biolabs) and RNA transcribed using T7 RNA polymerase. The RNA, diluted in the 1 X HBSS/glucose/DEAE-dextran mix, was used to transfect cell monolayers. A quarter of the mix was used to transfect HEp-2C cells in a $25cm^2$ flask and another quarter was used to transfect L20B cells in a $25cm^2$ flask. These were incubated in liquid culture at 35°C until CPE was observed. The rest of the RNA was used to make serial 10 fold dilutions in the 1 X HBSS/glucose/DEAE-dextran mix. HEp-2C and L20B cells in six well plates were inoculated with serial log dilutions, overlayed and incubated for three days. Plaques were picked following visualisation with neutral red before the plates were stained with naphthalene black. HEp-2C cells were used because they exert little selection pressure on poliovirus so that virus stocks would be representative of the original virus. L20B cells were also used to exploit the inability of changes in 2A to compensate for 5'NCR changes in these cells. Although poliovirus grows to slightly lower titres in L20B cells (Pipkin *et al.*, 1993) the selection pressure exerted by these cells to change in the 5'NCR could give rise to informative variants.

In chapter six, viruses were recovered from clones at least in duplicate to confirm the observed results. In contrast, the recovery of viruses described here was carried out once

from each clone. Controls were also used as for all transfections: one set of cells was mock transfected with the 1 X HBSS/glucose/DEAE-dextran mix and one set of cells was transfected with RNA made from a linearised clone of Leon/Lansing. In all cases, the mock transfected cells did not show CPE where as the Leon/Lansing transfected cells displayed CPE after 24-48 hours. Flasks of transfected cells were checked daily and frozen when complete CPE was detected. The supernatant from these flasks formed stocks of the resulting viruses. When no CPE was detected flasks were frozen after six days.

All HEp-2C derived virus stocks were checked for sequence at positions 483, 484, 514 and 528. This was carried out by RNA extraction, cDNA synthesis and PCR amplification with primers PCR F and PCR 9 (see chapter two). Purified PCR products were sequenced using primer αLL390 (see chapter two). Virus stocks were also assayed in BGM cells to measure the ts phenotype resulting from the mutations.

7.2.1 VIABILITY OF VIRUSES

Of all the combinations of bases at 483, 484, 514 and 528 that were tested, only one resulted in a non viable virus. This was LL/AG/GC. On assay in BGM cells at 35 °C, no plaques were obtained, even when the supernatant stock was used undiluted, from either HEp-2C or L20B cells. No PCR product could be obtained from RNA extracted from the supernatant stocks and positive controls were used to test the procedure. All other transfections were successful and the HEp-2C grown virus stocks were assayed in BGM cells. The viruses, LL/UA/GC, LL/AU/GA and LL/AU/AU were also assayed. Results can be seen in table 7.2 below.

VIRUS	Log_{10}(pfu at 35°C/pfu at 38°C) for BGM cells	Log_{10}(pfu at 35°C/pfu at 39°C) for BGM cells
Leon/Lansing AU/GU	0.5	0.8
LL/GC/CC	2.0	>2.0*
LL/CA/CG	2.2	3.1
LL/CC/CC	1.7	2.2
LL/CG/CA	2.5	4.3
LL/CA/UC	2.3	>2.3*
LL/AG/GC	-	-
LL/CC/UC	0.7	>1*
LL/AU/GC	1.2	2.5
LL/AU/AU	1	4.5

TABLE 7.2 Table to show ts phenotypes in BGM cells of viruses with differences at 483/528 and 484/514. Values shown are the log difference between titre at 35°C and the higher temperatures of 38°C and 39°C. For some viruses a finite value could not be obtained for the higher temperature as the titre of the virus was so low.

Observation of the domain V structure shown in fig 7.1 reveals that the potential for base pairs 483/528 and 484/514 relies fundamentally of the bases found at these positions. Clearly, either there will be a mismatch or a pair formed, leaving four possible combinations in each virus. They are: pair/mismatch; mismatch/pair; pair/pair and mismatch/mismatch. All naturally occurring viruses have the pair/mismatch version as found in the virulent, non ts Leon type virus. However, from the results in table 7.2 above, there does not appear to be a finite connection between the bases formed at this point and ts phenotype displayed.

PAIR/MISMATCH VIRUS	Log_{10}(pfu at 35°C/pfu at 38°C) for BGM cells	Log_{10}(pfu at 35°C/pfu at 39°C) for BGM cells
Leon/Lansing AU/GU	0.5	0.8
LL/CG/CA	2.5	4.3
LL/GC/CC	2.0	>2.0
LL/AU/GA	3.3	4.3

TABLE 7.3 Temperature sensitivity of viruses with a pair/mismatch genotype at bases 483/528 484/514 of domain V.

The viruses LL/CG/CA and LL/GC/CC should perhaps be discussed in the 'pair/pair' section below as the C in position 484 can form a base pair with the G at 527 at the end of the loop adjacent to stem (b) (see fig 7.3). This results in a domain V structure with an eight base stem (b), a thirteen base loop and no base in the hinge position. Calculated free energy values of these structures are quite low at -27.6 kcal/mol (*c.f.* -27.5 kcal/mol for Leon domain V) making them thermodynamically stable.

Folding of domain V could be disrupted to some extent as a result of different stacking of the loop sequence or because stem (b) is too long. Lengthening stem (b) has not been previously seen although a less ts variant of LLΔ483 had a stem (b) that was reduced from its parental six base pairs to five. This virus suggested a certain flexibility at this area of domain V provided by the 'hinge' base between stems (b) and (c) was more important than length of stem (b). These viruses without a 'hinge' base are ts and support this idea.

A virus with the ability to form a pair at 483/528 but no pair at 484/514 will have a domain V that has a seven base pair stem (b) and a fourteen base loop with a single base at the hinge (see fig 7.3). In the case of Leon/Lansing, this results in a viable, non ts, virulent virus, but in the case of LL/AU/GA, it results in a very ts virus. Comparison of these

viruses shows that the type of base pair at 483/528 has no effect as both have AU at this position. They also both have a G at the 484 hinge position. However the G in Leon/Lansing is able to form a weak base pair with the U at 514 whereas the G in LL/AU/GA has no potential base pairs. This does not support the hypothesis for the requirement of an unpaired 'hinge' base as LL/AU/GA would be predicted to be non ts. The only difference between these viruses is at position 528, the end base in the loop adjacent to stem (b). It is possible that a difference at this position interferes with the stacking or folding of the loop and causes the virus to be ts. This virus is however very ts and it would appear unlikely that a simple perturbation in the folding of the loop would bring about such a strong effect.

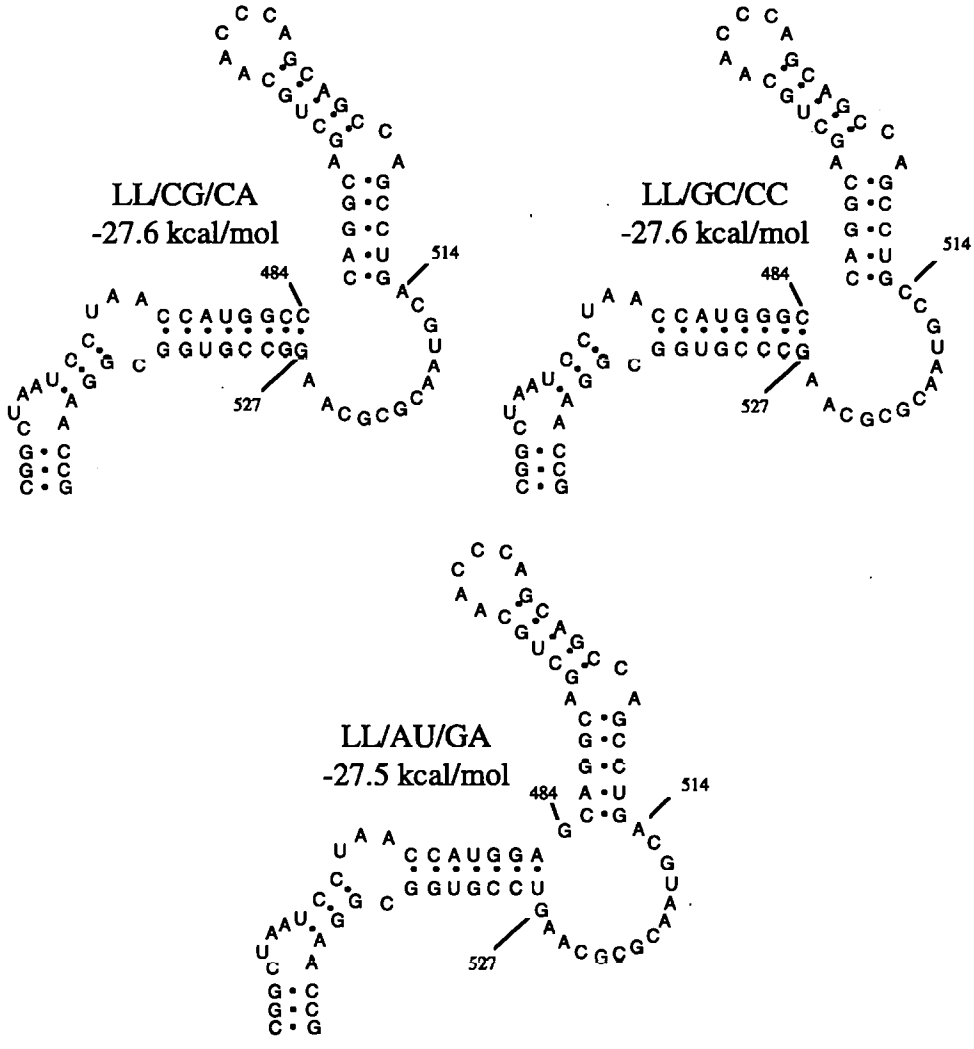


FIGURE 7.3 Probable domain V structures of viruses Leon, LL/CG/CA, LL/GC/CC and LL/AU/GA. The structures were calculated using 'foldRNA' and 'squiggles' GCG programs. Free energies are indicated.

MISMATCH/PAIR VIRUS	Log_{10}(pfu at 35°C/pfu at 38°C) for BGM cells	Log_{10}(pfu at 35°C/pfu at 39°C) for BGM cells
LL/CA/CG	2.2	3.1
LL/AG/GC	NON VIABLE	NON VIABLE

TABLE 7.4 Temperature sensitivity of viruses with a mismatch/pair genotype at bases 483/528 484/514 of domain V.

A mismatch/pair combination at this area of domain V would result in a structure with only six base pairs in stem (b), a fourteen base loop with a single base hinge and an increase to six base pairs in stem (c) (see fig 7.4). The mismatch/pair combination might be thought of as the 'opposite' of that found in Leon/Lansing but is such that a single 'hinge' base can still form. Previous results would predict that these viruses would be non ts when in actual fact they are very ts or non viable.

The requirement for flexibility at the 'hinge' base could feasibly be altered in the case of LL/AG/GC. The bases in the loop are such that a strong base pair could form between the G at 528 and the C at 514 in the large loop between stems (b) and (c). This would disrupt the secondary structure quite considerably and possibly promote the pairing of other bases in the loop which is thought to be unpaired according to analysis by strand specific reagents. Although there is potential for base pairing in this loop in the Leon domain V structure it does not occur so near to the 'hinge'. This would add further support to the idea that flexibility at this region is required.

In previous chapters the reduction of stem (b) does not appear to be very detrimental to the virus. The increase in base pairs in stem (c) however has not been addressed until now.

This could feasibly affect the way that the unpaired bases at the 'apex' of this domain V are presented to interact with either RNA in tertiary structure or with protein factors. Although there is some variation in this loop structure formed by bases 496-501 (see chapter four) this does not totally rule out interactions at this point.

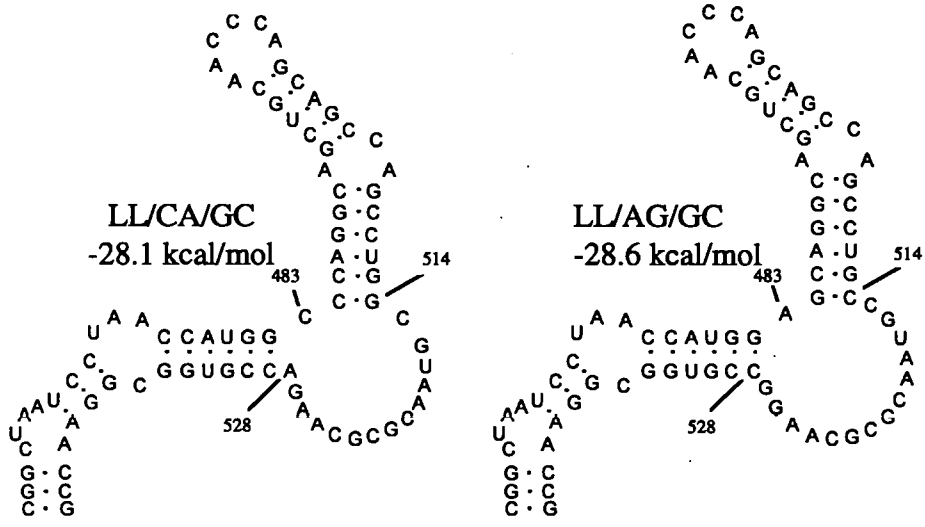


FIGURE 7.4 Probable domain V structures of viruses LL/CA/GC and LL/AG/GC as calculated using 'foldRNA' and 'squiggles' programs on GCG. The free energies of the structures are indicated.

7.5 PAIR/PAIR

PAIR/PAIR VIRUS	Log_{10}(pfu at 35°C/pfu at 38°C) for BGM cells	Log_{10}(pfu at 35°C/pfu at 39°C) for BGM cells
LL/AU/GC	1.2	2.5
LL/AU/AU	1	4.5

TABLE 7.5 Temperature sensitivity of viruses with a pair/pair genotype at bases 483/528 484/514 of domain V.

With base pairs at both 483/528 and 484/514 able to form, the domain V structure of these viruses could be quite constrained at this area. Stem (b) would have seven base pairs, as found in the Leon type virus, stem (c) would be increased to six base pairs and the loop would be decreased to thirteen bases (see fig 7.5). This loop also has potential base pairing between the G at 527 and the C at 515, equivalent to the potential base pairing discussed

for LL/AG/GC. This extra base pair, combined with the lack of 'hinge' base would again alter the folding and flexibility of domain V at this area. The calculated free energy of these structures are relatively high. This contrasts with the low free energy values calculated for LL/CG/CA and LL/GC/CC which also have no 'hinge' base and are similarly ts. Again the lack of a 'hinge' would appear to be a major determinant of a ts phenotype.

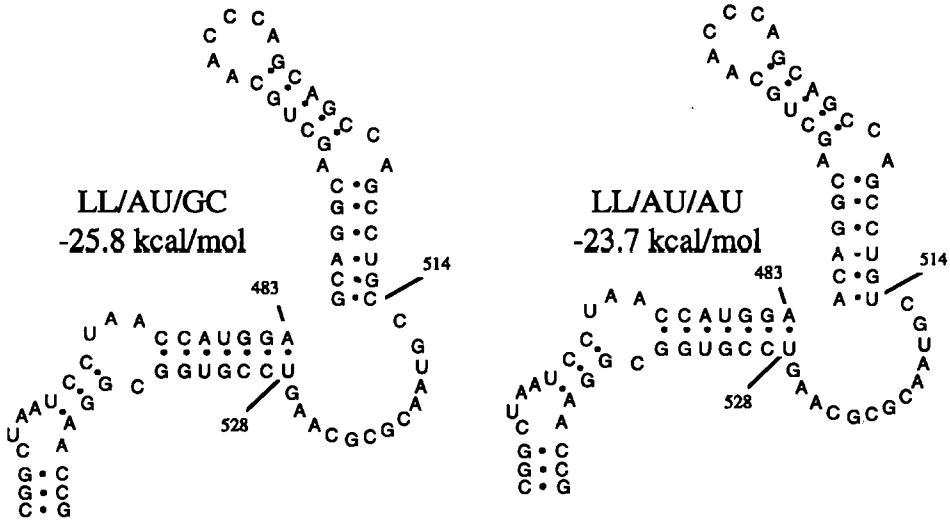


FIGURE 7.5 Probable domain V structures of viruses LL/AU/GC and LL/AU/AU calculated using 'foldRNA' and 'squiggles' GCG programs. Free energies are indicated.

7.6 MISMATCH/MISMATCH

MISMATCH/MISMATCH VIRUS	Log_{10}(pfu at 35°C/pfu at 38°C) for BGM cells	Log_{10}(pfu at 35°C/pfu at 38°C) for BGM cells
LL/CC/UC	0.7	>1
LL/CC/CC	1.7	2.2
LL/CA/UC	2.3	>2.3

TABLE 7.6 Temperature sensitivity of viruses with a mismatch/mismatch genotype at bases 483/528 484/514 of domain V.

These viruses have no ability to base pair at the positions being studied here. Therefore they have domain V structures that have a reduction to six base pairs in stem (b), two bases

in the hinge position and fifteen bases in the loop (see fig 7.6). As previously discussed extra bases in the loop and hinge could affect the stacking of bases in the loop and the folding of the entire structure. This could conceivably contribute to the ts phenotype of a virus. However from the results in table 7.6 there is no direct link as the phenotype these viruses express varies. However LL/CC/CU may appear to be relatively non ts due to revertants arising during the assay.

The viruses LL/CC/CC and LL/CC/UC are also capable of forming an alternative structure involving stem (b) and these are shown in fig 7.6. Both viruses have a C at 483 and this enables base pairs to form between 482/528 and 483/527, reducing the loop to thirteen bases, whilst the C at 529 in stem (b) is forced to unpair. A single base is then left in the 'hinge' position. The unpaired base in stem (b) could conceivably alter the folding of the stem and the entire domain V structure, contributing to the ts phenotype of the virus. In both cases the free energy is decreased by 1 kcal/mol, making this the more energetically favourable form but with such a small energy difference it is possible that both structures are present in the virus population. Two bases in the 'hinge' position would therefore appear to be as detrimental to the virus as the absence of a 'hinge' base.

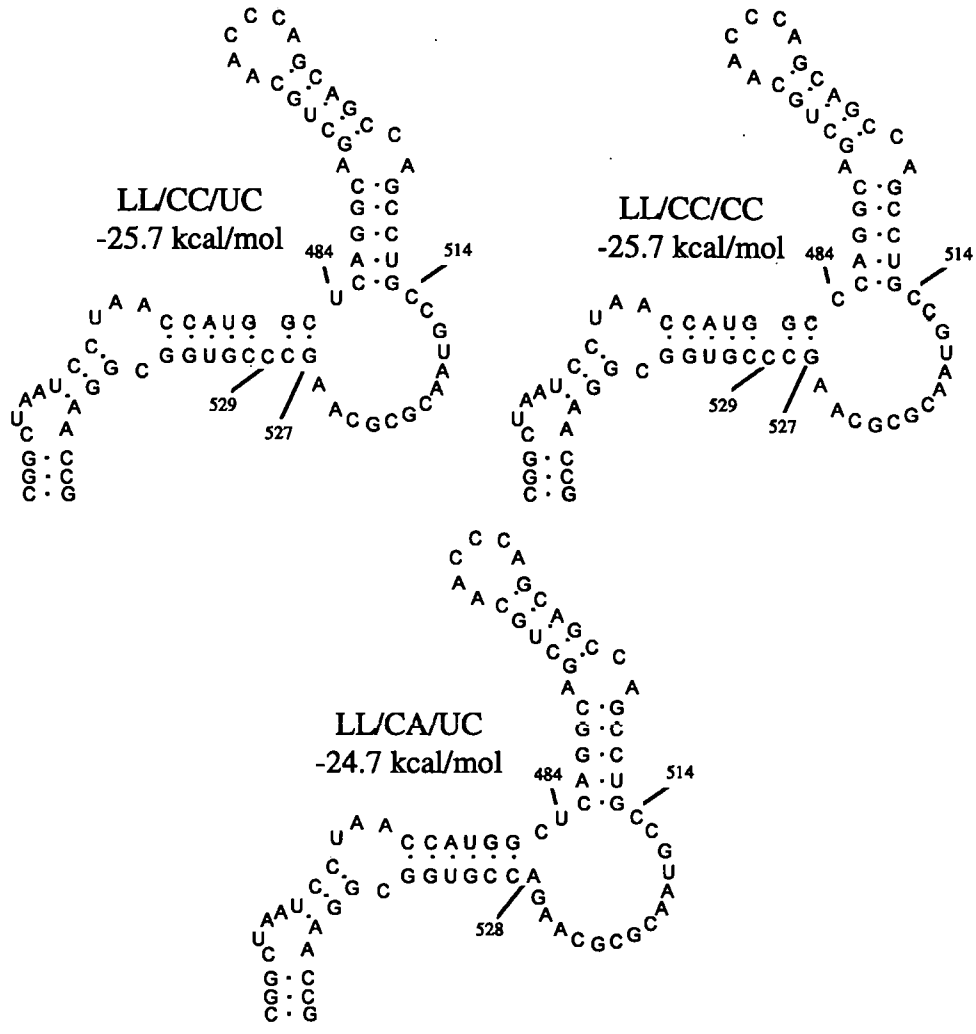


FIGURE 7.6 Probable domain V structures of viruses LL/CC/UC, LL/CC/CC and LL/CA/UC calculated using 'foldRNA' and 'squiggles' GCG programs. Free energies are indicated.

7.7 DISCUSSION

The results presented in this chapter are somewhat confusing. While on the whole the requirement for a finite flexibility from a single base in the 'hinge' position at 484 is upheld, not all the phenotypes of the viruses generated here can be totally rationalised in this way.

A mismatch/mismatch combination would result in two bases forming the 'hinge' and a pair/pair combination leaves no base at all in this position. All of these viruses are ts and would support the requirement for a single 'hinge' base except LL/CC/UC. The phenotype

of this virus cannot therefore be explained in terms of 'hinge' requirements. As sequencing confirmed the bases present it is possible that revertants may have arisen during the assay making the virus appear non ts. As the initial growth of the virus in HEp-2C cells did not appear to produce revertants so easily this would seem to be unlikely. Analysis of plaques from this virus would establish if this was true.

Some of the ts phenotypes could be a result of the potential base pairs in the thirteen base loop between stems (b) and (c). Although some potential base pairings exist in the Leon domain V loop these are not thought to form (see previous chapters) and therefore would appear to be detrimental to the functional structure. However not all of the viruses have this potential and remain ts. LL/AG/GC would be expected to have a single 'hinge' base and therefore be non ts but instead it is non viable. Both LL/AU/GA and LL/CA/CG would also be predicted to have a single 'hinge' base and little potential base pairing in the loop yet are still ts. The phenotypes of these viruses do not support the single 'hinge' base theory and cannot be rationalised in terms of requirements of functional structure at the moment unless the folding of the unpaired bases in the loop is considerably altered by a slight change in sequence or length.

The effects that the changes at the four bases around the loop exert are complex and in some cases allow alternative structures to form. Ideally other mutations should be introduced into the virus to further analyse the requirement of bases at this position. As other features of domain V were also altered in these clones it would be useful to singularly analyse the effects of these changes on the virus. Future clones to be made would include a deletion of only the 'hinge' base at 484, deletion or addition of bases to the thirteen base loop, substitutions of bases in the loop and adding or deleting base pairs in stem (b) and (c) (LLΔ483/527 already exists). The addition of an extra base into the 'hinge' without

forming extra base pairs would also be informative but unfortunately this appears to be virtually impossible to achieve. Reversion of Sabin 2 always leaves a G at 481 because any other base in this position would result in the formation of other base pairs. However it is possible that inserting a C after the G in position 484 would fulfil this although the free energy of this is predicted to be high at -18.7 kcal/mol (*c.f.* -27.5 kcal/mol for Leon).

The results here show that the simple 'hinge' theory does not fully account for the sequence requirements of domain V in the region between stems (b) and (c). They also do not suggest an obvious alternative model. Analysis of revertants or variants of these viruses from plaque selection will undoubtedly give more insight as to the functional structure required. It might be expected that a single 'hinge' would be formed wherever possible and that potential base pairing in the loop would be removed if the discussions above have been correct. In addition revertants may also reveal why LL/AU/GA and LL/CA/CG are ts. The reconstruction of LL/AG/GC should also be repeated to ensure the virus is truly non viable.

CHAPTER EIGHT

GENERAL DISCUSSION

8.1 GENERAL DISCUSSION

The positive strand polioviruses have an unusually long 5'NCR with extensive secondary structure that is essential for replication and translation of the genomic RNA. Part of the secondary structure forms an internal ribosome entry site (IRES) which is able to steer the ribosomal initiation complex directly onto the RNA near to the initiating AUG codon. In contrast, most cellular RNA molecules have a cap structure at their 5' ends and initiation of translation from them occurs by a scanning, cap-dependent mechanism. Polioviruses are able to inhibit cap-dependent translation and force an infected cell to preferentially synthesise viral proteins by the alternative cap-independent method. This is achieved, at least in part, by the cleavage of the initiation factor eIF4G by the viral protease 2A, which effectively removes the cap binding protein, eIF4E, from the initiating complex, eIF4F. Efficient initiation of viral translation requires an intact stem-loop secondary structure of the IRES along with canonical cellular factors but is also enhanced by many other protein factors.

The IRES is composed of domains II, IV, V and VI of the 5'NCR. Mutations in domain V which alter secondary structure confer a ts phenotype and reduce viral translation efficiency. Coding changes in poliovirus protease 2A are able to compensate *in vitro* for such mutations in the 5'NCR. This phenomenon was first reported in Macadam *et al.*, (1994) and consolidated by the results described in chapters three and five of this thesis, based on analysis of revertants and site directed mutants. The protease 2A is known to have several functions in the lifecycle of the virus, acting mainly as a protease both in polyprotein processing and in host cell translation shut off. Involvement of 2A as a transactivator of poliovirus IRES driven translation was first reported by Hambidge & Sarnow (1992). This function was also later found to be a property of the human rhinovirus 2A proteinase acting with the cognate IRES (Liebig *et al.*, 1993).

Perhaps surprisingly a large number of different coding changes found along the entire length of 2A are able to compensate for a ts phenotype in BGM cells. The protease is fairly highly conserved among polioviruses. Limited variation of 2A was confirmed from sequence analysis of eight wild type polioviruses which were isolated over a number of years from a variety of countries. No X-ray crystallographic information exists for the structure of 2A, merely a putative α–carbon chain model of the protease based on sequence alignments with serine proteases. On the basis of the model, changes appear to be on the external surface of the protein, suggesting they affect interaction with other proteins or RNA molecules. However as the protein is relatively small this finding is not necessarily significant. As a result the structural consequence of the coding changes is difficult to resolve. Comparisons of picornaviral 2A sequences revealed variation specific to a genera and to viruses within a genus, in agreement with current published data. This was thought to reflect adaptation to different growth conditions.

The discovery that the compensatory mutations were cell specific indicates that this property of 2A requires cellular factors and that 2A probably does not interact directly with the 5'NCR. Interaction could involve general protein-protein association or factor cleavage by the proteolytic activity of 2A. The precise mechanism by which the protein is acting in this function is still unknown. It was suggested that the transactivating function of 2A existed so as to extend the range of cell types that polioviruses can infect (Hambidge & Sarnow 1992). Indeed changes in 2A, along with changes in VP1, were associated with mouse neurovirulence of the adapted mouse strain LS-a (Lu *et al.*, 1994). Interestingly two of the three changes observed by Lu *et al.*, (1994), at residues 25 and 70, were found to be at positions where compensating mutations were identified. The change at 70 was the same as that found in a Leon/Lansing non ts revertant reported here. It was proposed

therefore that the protease acts to enhance translation rather than as a fundamentally required factor, at least *in vitro*.

The transactivational activity of 2A was reported to be related to the protease activity (Ventoso & Carrasco 1995). Mutant 2A proteins with changes at residue 88 which were unable to cleave 3CD also had a reduced ability to enhance translation from a reporter gene under control of a poliovirus IRES in HeLa cells. Cleavage of eIF4G by the mutant protein was however not affected, indicating that the substrate specificity of 2A can be differentially altered. Proteolytic activity was also reported to be a requirement for full enhancement of rhinovirus translation by rhinovirus 2A (Ziegler *et al.*, 1995a). Reduced translational enhancement activity in rabbit reticulocytes was shown to result from either inhibiting 2A proteolysis with elastatinal or mutating the catalytic cysteine to serine.

However there is no evidence of the compensatory 2A mutations described in this thesis as having any effect on proteolytic activity. Although it is difficult to position these changes on the putative 2A model it is clear that most are not found at or near the catalytic triad. Indeed none of the changes were found to alter the highly conserved GDCGG proteolytic motif. In addition, protein translation experiments in Macadam *et al.*, (1994) showed no effect on the shut off of host cell protein synthesis. In fact 2A mutations were found to act at a stage after the inhibition of host cell translation was achieved. This meant that 2A had a role in cap-independent translation other than cleavage of the initiation factor.

It is possible however that the changes to 2A alter the equilibrium of the reaction between the protease, its substrate and cleavage products, *e.g.* releasing more cleaved eIF4G for initiation of cap-dependent translation. This would provide some link of the changes to the proteolytic activity of 2A. Only the C-terminus of cleaved eIF4G is required for the

translation of uncapped RNA, with or without an IRES element (Ohlmann *et al.*, 1995; Ohlmann *et al.*, 1996; Pestova *et al.*, 1996) and eIF4G was suggested to be a factor involved in translation stimulation by Ziegler *et al.*, (1995a & 1995b). However eIF4G is unlikely to be the factor proposed to interact with 2A because there is no evidence of 2A being inefficient in its role of inhibiting cap-dependent translation. Indeed, proteolysis of eIF4G was observed even when poliovirus replication was severely inhibited so that only a low level of viral proteins will be present (Bonneau & Sonenberg 1987a). There is also no evidence for a fundamental factor such as eIF4G being absent or sufficiently different in either monkey CNS or L20B cells. Finally, the compensatory role of the 2A changes discussed here is an *in vitro,* cell specific observation and it is unlikely that the function observed involves a fundamental factor that is present in all cells.

Polioviral 2A proteinase is also reported to be required for replication. A bicistronic virus was constructed in which poliovirus 2A was replaced by an EMCV IRES between P1 and P2, removing the autocatalytic activity requirement of 2A (Molla *et al.*, 1993b). This virus was found to be non viable and negative results from RT-PCR lead to the conclusion that replication had been abolished. Similarly, site-directed mutation in 2A in full length clones were found to reduce RNA levels in transfected cells as shown by Northern blots (Yu *et al.*, 1995). However a reduction in RNA could equally be the result of a reduction in translation limiting viral proteins essential for replication. Results from protein labelling experiments in which infected cell sheets were incubated at the permissive temperature of 35°C prior to labelling at elevated temperatures provided no evidence that the compensatory changes in 2A discussed here affect replication rates (personal communication, A. J. Macadam). RNA dot blots of cytoplasmic lysates of infected cells also show no difference in RNA levels.

It is possible that the coding changes in 2A act so as to diminish a binding function of the protease. A binding efficiency could easily be reduced by any number of amino acid changes. The unknown protein factor, once released can then act to enhance initiation of translation, either binding to the disrupted RNA or other protein factors and then directing the ribosomal machinery onto the IRES. This binding function of 2A could possibly be an additional function to the transactivating function previously reported as no link with proteolytic activity was found. The protease could fundamentally enhance translation *in vivo* as well as compensating for translation inefficiencies. Published work involved translation in cell culture, rabbit reticulocyte lysates and nuclease treated cytoplasmic extracts. In contrast the cell specific function of 2A described in this thesis was not identified using a cell free system, in which 2A may perform differently, and this would have given confusing results.

The binding function discussed here could itself be temperature dependent. This would mean that for a ts virus, the efficiency of translation would rise again at a higher temperature as the factor was released. It is possible that this is never observed because at the higher temperature other functions of the virus or host cell are disrupted.

Site directed mutants with different codons for residue 79 of 2A indicate that not all coding changes are able to equally compensate a disrupted domain V. Indeed not all domain V disruptions can be compensated for by 2A mutations. However, the changes introduced into residue 79 also suggest that RNA sequence has some role in the compensatory function. From initial evidence a protein rather than RNA function was suggested as all non ts revertants were found to have coding changes. The lack of any silent changes still suggests that the protein has some compensatory role instead of the RNA but this is an issue that needs to be addressed in the future.

Changes in 2A appear to pose no threat to the monitoring of vaccine strains as they have no effect on monkey neurovirulence. Changes in 2A arising during vaccine manufacture are unlikely to affect test results and in fact may stabilise attenuating mutations. One of the compensatory changes in 2A, a change of serine to threonine at residue 134 is also found in Sabin 1. It is possible that this change is able to enhance initiation of translation from the Sabin 1 5'NCR which is disrupted by the change at 480 in domain V. This would remove some pressure on the virus to revert in domain V and help to explain why reversion of 480 is slower than that of *e.g.* 472 U in Sabin 3 (Minor & Dunn 1988).

The discovery of these 2A changes does however demonstrate the unpredictability of a virus. Live attenuated vaccines will always need to monitored to ensure they will not cause adverse reactions to the vaccinee. A monkey neurovirulence test is currently the only means of assaying the vaccine's potential to cause human disease at the moment. The test relies on the assumption that monkey neurovirulence is similar enough to human neurovirulence for a significant result. A transgenic mouse neurovirulence test could be feasible in the same way. Unfortunately molecular biology assays such as the MAPREC (mutant analysis by PCR and restriction enzyme cleavage) test (Chumakov *et al.*, 1991) may never be able to replace the use of animals. Although the major determinant of Sabin 3 attenuation is a mutation in domain V which can easily be detected by MAPREC, such an assay only tests the vaccine at a single site and will not detect spontaneous changes occurring elsewhere.

The observation that changes to 2A were unable to compensate for domain V disruptions in L20B cells provides an opportunity to identify the enhancing factor that is proposed to interact with 2A. Expression of BGM cDNA libraries in L20B cells using a modified CELICS technique may allow identification if the factor is totally absent from L20B cells

or in limiting amounts. The CELICS method involves a β-galactosidase antibody conjugate which gives a blue colour in the presence of X-Gal and was used to identify the receptor for echovirus 7 (Ward et al., 1994).

Models of 5'NCR secondary structure were initially based on computer predictions, biochemical probing and supported by mutagenesis and revertant analysis. However, little is known about the region's tertiary structure which is probably formed through sequence specific interactions of unpaired loops or stacking together of helical stems. The discovery that compensatory 2A coding changes were cell specific provided a useful system for the study of domain V and 5'NCR structure where reversion of site-directed mutants would not be complicated by changes in 2A.

Using this system (L20B cells), revertants of Sabin 2 were only found with an A to G change at 481 and no second site mutation appears to be able to compensate for this disruption to the secondary structure. Also using this system, variants of LLΔ472 and site directed mutants revealed that the three base loop made up of bases 474-476 (type 3 numbering) could accommodate sequence variation but not reductions in loop length. This suggested that the unpaired bases in the loop did not interact with either RNA or proteins but rather provide a spacer of specific length between base paired stems. In addition, variants of LLΔ483 implied a single 'hinge' base at 484 (type 3 numbering) with as little potential base pairing was required to give a specific flexibility between stems (b) and (c). Analysis of site directed mutants produced inconclusive results and the exact requirements in this region of domain V are still unknown. Other published work on the requirements of bases in the 474-476 loop and 484 'hinge' base have relied on crude insertions of linker sequences into

domain V and revertants were only found to have re-formed some base pairings in stems (Haller & Semler 1992; Haller *et al.*, 1996).

Although no evidence for tertiary structure was obtained from reversion of ts viruses in L20B cells thus far, future experiments may help to pinpoint RNA/RNA interactions. Unpaired bases in loops are ideal candidates for this although sequence variation does indicate where this is unlikely. Comparison of sequencing evidence shows that the large loop between stems (b) and (c) is highly conserved. Mutations that alter the length or sequence of this loop were also implicated in the ts phenotypes of site directed mutant viruses. However the sequence of the three base loop between stems (a) and (b) is also highly conserved yet sequence variation has only a limited effect on the virus. Further sequence comparisons of other domains would also identify potentially interacting loops. It is also possible that the six domains exist as individual entities until the ribosomal complex forms for initiation of translation. However close proximity of the domains would suggest that some kind of tertiary structure will form, perhaps altering to some extent when interacting with translational proteins.

The WHO's initiative to eradicate poliomyelitis and polioviruses is currently under debate and programs of when to stop vaccinations are being discussed. It is likely that eradication will be successful in the near future and this will mean that research on the virus will eventually be stopped. It is possible therefore that all polioviruses will be destroyed before some of the questions that have arisen from this thesis are answered. Issues may instead be addressed with other viruses in the picornavirus family such as enteroviruses and rhinoviruses.

APPENDIX

APPENDIX 1

```
                3377                                                 3426
P1.1            .......... .......... .......... .......... ..........
P1.2            .......... .......... .......... .......... ..........
P1.3            .......... .......... .......... .......... ..........
P2.1            .......... .......... .......... --a------- cg--a-----
P2.2            .......... .......... .......... .......... ..........
P2.3            .......... .......... .......... .......... ..........
P3.1            .......... .......... .......... .......... ..........
P3.2            .......... .......... .......... .......... ..........
P3.3            .......... .......... .......... .......... ..........
P3/LEON         ggcuuugggc aucagaauaa agcuguguac --u------- ac--g-----
CONSENSUS       ---------- ---------- ---------- AC-GCUGGUU --AA-AUCUG

                3427                                                 3476
P1.1            .......... .......... .......... .......... ..........
P1.2            .......... .......... .......... .......... ..........
P1.3            .......... .......... .......... .......... ..........
P2.1            ---------- --a----ag- --------c- g--------- --------u-
P2.2            .......... ...-----a- -a-------- ugc------- -----ug---
P2.3            .......... .......... .......... .......... ..........
P3.1            .......... .......... .......... .......... ..........
P3.2            .......... .......... .......... .......... ..........
P3.3            .......... .......... .......... .......... ..........
P3/LEON         ---------- --c-----ua ----g--u-u a------a-u --a-------
CONSENSUS       CAACUACCAU CU-GCCAC-C AGGAAGACUC -CAAAAUGCA GUGAGCAUCA

                3477                                                 3526
P1.1            .......... .......... .......... .......... ..........
P1.2            .......... .......... .......... .......... ..........
P1.3            .......... .......... .......... .......... ..........
P2.1            ---------- ---------a a--------- -c---agc-c ---------u
P2.2            ---------- ---u------ ---------- ---------- g--------u
P2.3            .......... .......... .......... .......... ........--c
P3.1            .......... .......... .......... .......... ..........
P3.2            .......... .......... .......... .......... ..........
P3.3            .......... .......... .......... .......... ..........
P3/LEON         ---------- ------c--- --ugu----- ----a----- ---u--c--c
CONSENSUS       UGUGGAAUAG AGACCUAUUG GUCACUGAAU CAAAGGCUCA AGGAACAGA-

                3527                                                 3576
P1.1            .......... .......... .......... .......... ..........
P1.2            .......... .......... .......... .......... ...-g-a---
P1.3            .......... .......... .......... ......--c- ----c-----
P2.1            --g------- -a-------- ------u--- --u-----c- ----a-----
P2.2            --------c- -a-----u-- ----g-a--a --------u- ----a--u--
P2.3            --g-----u- -g----g--- ---c--g--- --------c- ----a-----
P3.1            .......... .......... .......... .......... ..........
P3.2            .......... .......... .......... .......... ..........
P3.3            .......... .......... .......--- -----u--u- ----g--u--
P3/LEON         -----a--a- -g-----u-- c---g-a--g --------u- ----g-----
CONSENSUS       UCAAUCGCGA G-UGCAACUG UAAUAC-GGU GUGUACUA-U GUGA-UCCAG
```

```
                3577                                                3626
P1.1            .......... .......... .......... ..-------- ----------
P1.2            ---------- ---------- -------u-- ---u------ ----------
P1.3            ---------- ---------- -------a-- c--g-----u ----------
P2.1            ---------- ---------- -------u-- ---u------ -----u----
P2.2            a--------- ---------- ----u--c-- g-----u--- ----------
P2.3            ---------u --u-----u- -------u-- g--------- --a-------
P3.1            .......... .......... ....ugga-- ---g------ --a-------
P3.2            .......... .......... .......... .......... ..........
P3.3            a-----g--- ---------- -c-a-g-a-- c--a-----u ----------
P3/LEON         a--g------ -----u--g- -g--ug-g-- ---------u ----------
CONSENSUS       GAGAAAAUAC UACCCAGUCU CUUUCAU-GG ACCCACCUUC CAGUACAUGG

                3627                                                3676
P1.1            ----c----- u--u------ --------c- ----c--u-- ---c------
P1.2            ---------- u--------- -------g-- ---------- ---c--a---
P1.3            -a-------- g-----u--- -----a---- -g--c----- ----------
P2.1            ---u------ u--------- ---------- ---------- ---c--a---
P2.2            ---------- a-----u--- ---------- ---------- ------a---
P2.3            -a--c----- a--u-----g --u--a---- ---------- ---u-----u
P3.1            ----u--c-- c--------- --------c- ----c----- ------c-c-
P3.2            .......... .......... .......... .......... ..........
P3.3            -a-------- c-----u--- ---------- -g-------- -u-g-----c
P3/LEON         ----u----- c--------- --u--a--c- ----c----- -u----c---
CONSENSUS       AGGCAAAUGA -UACUACCCA GCCAGGUAUC AAUCACACAU GCUAAUUGGG

                3677                                                3726
P1.1            -----a--.. .......... .......... ....-au--- ----------
P1.2            --c-----c- ---------- ---------c -----u---- ----------
P1.3            -----a---- ---------- ---c------ ---------- -g--c--g--
P2.1            --c-----c- ---------- ---------- -----u---- ----------
P2.2            --------c- ---------- ---a------ ---------- -g--------
P2.3            -----g---- -------g-- ---------- --c--a--u- ----------
P3.1            u-c------- -------g-- ---------- --c------- ----c-----
P3.2            .......... .......... .......... ...-----g- ----c-----
P3.3            -----g---- -------g-- ---c-----a --g--u---- -g--------
P3/LEON         --c--c---- ---------- ---c------ --------u- -g--------
CONSENSUS       CAUGGUUUUG CAUCACCAGG -GAUUGUGGU GGUAUCCUCA GAUGUCAACA

                3727                                                3776
P1.1            ---------- ---------- -u--g----- ------g--g --a-------
P1.2            ---------- ---------- -u--g----- ------g--g --a-------
P1.3            c--c------ --g--ug--- -a-----a-- -------c-- --c--u----
P2.1            ---------- ---------- -a-------- ------g--g --a---c-u.
P2.2            ---------- ---------- -u-------- g------c-- --g-----u-
P2.3            c--------- ---------- -u-------- g---------g --u--a--u
P3.1            c--a-----u ------g-a- -a-----c-- g--g--u--- --g--u----
P3.2            c--a-----u ------g-a- -a-----c-- g--g--u--- --c--u----
P3.3            ---a------ --c--u---- -u--a--a-- -------c-c -au--g----
P3/LEON         ---c-----c -----cg-g- -a-------- ---g--a--- --c--a----
CONSENSUS       UGGGGU-AUA GGAAUAAUCA C-GCUGGUGG AGAAGGCUUA GU-GCCUUCU
```

```
                3777                                                  3823
P1.1            ---------- ------u--- --u------- ----.....  .......
P1.2            ---------- ---.....  ..........  ..........  .......
P1.3            ----u--c-- g-----a--c --c------- ----g-----  ---a--a
P2.1            ..........  ..........  ..........  ..........  .......
P2.2            -------c--  ---u--u---  --a-------  ----a-----  ---g--a
P2.3            ----------  ----------  --u-----g-  -g--a--u--  ---g--g
P3.1            -c---u-a..  ..........  ..........  ----a--u--  -c-a--g
P3.2            -c-----a--  g--------c  --c-----g-  ----a-----  ---a--g
P3.3            -ga-------  ------c---  --c-------  -g--g-----  ---a--a
P3/LEON         -u-----a--  g---------  --u--c--g-  ----g-----  ---g--g
CONSENSUS       CAGACAUUAG  AGACCUGUAU  GC-UAUGAAG  AAGA-GCCAU  GGA-CA-
```

APPENDIX 1 A comparison of nucleotide sequences of the region coding for the protease 2A obtained from nine wild type viruses. Alignment was calculated using GCG 'pileup' and 'pretty' programs. Conserved residues are shown as dashes (-) and differences to the consensus are in lower case. Gaps in the sequence are shown as full stops (.) where the sequence is incomplete. Numbering is from type 3 Leon.

APPENDIX 2

```
                3377                                                    3426
P1/Sabin        --a--c--a- -------c-- ---g------ -----a---- ----a--u--
P1/Mahoney      --a--c--a- -------c-- ---g------ -----a---- ----a--u--
P2/Sabin        --a-----a- -------c-- ---------- --a-----c- ----a--u--
P2/Lansing      --u-----c- ---------- g--a------ --g--a---- ----a--u--
P2/W2           --u-----c- ---------- g--a------ --g--a---- ----a--u--
P3/Leon         ---------- -u--g----- ---------- ---------- ----g--c--
P3/Sabin        ---------- -u--g----- ---------- ---------- ----g--c--
P3/Finland      --g-----u- ---------- ------u--- --a--g---- -u--g--c--
P3/Leon 119     ---------- -u--g----- ---------- ---------- ----g--c--
CONSENSUS       GGCUUUGGGC ACCAAAAUAA AGCUGUGUAC ACUGCUGGUU ACAAGAUCUG

                3427                                                    3476
P1/Sabin        ---------u u-g------- ---------- g-----c--a --g-a-g---
P1/Mahoney      ---------c u-g------- ----u----- g-----c--a --g-a-g---
P2/Sabin        ---u-----c --a--u--a- -a-----c-- g--------c --g--ug---
P2/Lansing      ---u-----c --------c- -------c-- ---------g ----a---u-
P2/W2           ---u-----c --------c- -------c-- ---------g a---a---u-
P3/Leon         ---------u ---------a ----g----- ---------- ----------
P3/Sabin        ---------u ---------a ----g----- ---------- ----------
P3/Finland      ---u-----c -----u--a- -a------c- ---g-----a --g---g---
P3/Leon 119     ---------u ---------a ----g----- ---------- ----------
CONSENSUS       CAACUACCAC CUCGCCACUC AGGAAGAUUU ACAAAAUGCU GUAAGCAUCA

                3477                                                    3526
P1/Sabin        ---------- ---------a --caca---- ---g---c-- g--------u
P1/Mahoney      ------g--- ---------a --caca---- ---g---c-- g--------u
P2/Sabin        -------c-- ---------a --g- c---- ---gg--c-u u---------
P2/Lansing      ------u--- ------u--a --a--g---- -c-----c-- -----ua---
P2/W2           ------u--- ------u--a --a--g---- -c-----c-- -----ua---
P3/Leon         ---------- ---------g ---------- -------u-- ---u------
P3/Sabin        ---------- ---------g ---------- -------u-- ---u------
P3/Finland      ---------- g--u---c-g --gaca---- -------u-- g----ua--u
P3/Leon 119     ---------- ---------g ---------- -------u-- ---u------
CONSENSUS       UGUGGAAUAG AGACCUCUUA GUUGUUGAAU CAAAAGCCCA AGGCACCGAC

                3527                                                    3576
P1/Sabin        -----c---- ---------- ---------- --------c- -c-----u--
P1/Mahoney      -----c---- ---------- ---------- --------c- -c-----u--
P2/Sabin        --g--c---- ------gc-- u---a-g--u --------c- ----a-----
P2/Lansing      -----u--u- -a--u--c-- -c--a-u--a --------c- ----a-----
P2/W2           -----u--u- -a--u--c-- -c--a-u--a --------c- ----a-----
P3/Leon         -----a---- ---------- ---u------ --------u- ----------
P3/Sabin        -----a---- ---------- ---u--g--- --------u- ----------
P3/Finland      --u--c--c- -a-------- --g-a-u--c -----u--u- ----------
P3/Leon 119     -----a---- ---------- ---u------ --------u- ----------
CONSENSUS       UCAAU-GCAA GGUGCAAUUG CAACGCAGGG GUGUACUACU GUGAGUCCAG
```

```
                3577                                                 3626
P1/Sabin        ---------- -----a--a- -c--c----- c--a--g--- --g-------
P1/Mahoney      ---------- -----a--a- -c--c----- c--a--g--- --g-------
P2/Sabin        g--a-----u --u--a--u- -u--ca---- g--------- ----------
P2/Lansing      g-----g--- -----g--c- -u---ac--- c--------u --g-------
P2/W2           g-----g--- -----g--c- -u-----c-- c--------u --g-------
P3/Leon         ---------- ---------- -------g-- a--------- ----------
P3/Sabin        ---------- ---------- -------g-- a--------- ----------
P3/Finland      ---u-gg--u -------u-- -c--c----- u--a--u--- ----------
P3/Leon 119     ---------- ---------- -------g-- a--------- ----------
CONSENSUS       AAGGAAAUAC UACCCUGUGU CGUUUGUUGG -CCCACCUUC CAAUACAUGG

                3627                                                 3676
P1/Sabin        --------a- ---u------ -----g---- -g-----u-- -c-c--u--c
P1/Mahoney      --------a- ---u------ -u---g---- -g-----u-- -c-c--u--c
P2/Sabin        -a--c----- a--u-----g --------u- ----a----- -c-u--u--u
P2/Lansing      -a--a----- g-----u--- --cc------ ---------- ------u--c
P2/W2           -a--a----- g-----u--- --cc------ ---------- ------u--c
P3/Leon         ---------- ---------- ---------- ---------- ------c--g
P3/Sabin        ---------- ---------- ---------- ---------- ------c--g
P3/Finland      -a--a----- u--------- ---c------ -g-------- -c----c--u
P3/Leon 119     ---------- ---------- ---------- ---------- ------c--g
CONSENSUS       AGGCUAAUGA CUACUACCCA GCUAGCUACC AAUCCCACAU GUUAAUUGG-

                3677                                                 3726
P1/Sabin        --u--a--c- ----u----- g--u------ --c--a--c- -a-----c--
P1/Mahoney      --u--a--c- ----u----- g--u------ --c--a--c- -a-----c--
P2/Sabin        --u--g---- -------g-- u--u------ --c--a--u- -a--------
P2/Lansing      --u--u---- ----u----- g--------- --g--u--c- ----c-----
P2/W2           --u--u---- ----u----- g--------- --g--u--c- ----c-----
P3/Leon         --c------- -c-------- u--------- --------u- ----------
P3/Sabin        --c------- -c-------- u--------- --------u- ----------
P3/Finland      --c--u--c- -------u-- g--u-----c --------c- -a--c--g--
P3/Leon 119     --c------- -c-------- u--------- --------u- ----------
CONSENSUS       CAUGGCUUUG CAUCACCAGG GGACUGUGGU GGUAUCCUCA GGUGUCAACA

                3727                                                 3767
P1/Sabin        c--g--g--a --g-----u- -u-------- ---a--g--g --u-----ua
P1/Mahoney      c--g--g--a --g-----u- -u-------- c--a--g--g --u-----u-
P2/Sabin        c--g--g--a -----a--c- -u-------- g--a--c--g --u-----u-
P2/Lansing      ---a--a--u --------u- -------a-- ---a--cc-- -----u----
P2/W2           ---a--a--u --------u- -------a-- ---a--cc-- -----u----
P3/Leon         ------c--c ------g-g- ---------- ---g--a--- ----------
P3/Sabin        ------c--c ------g-g- ---------- ---g--a--- ----------
P3/Finland      ---u--g--a --c-----c- -g-----a-- ---g--cc-g --g--u----
P3/Leon 119     ------c--c ------g-g- ---------- ---g--a--- ----------
CONSENSUS       UGGCGU-AU- GGAAUCAU-A CAGCUGGUGG AGAAGG-UUA GUCGCAUUCU
```

```
                3777                                                  3823
P1/Sabin        -a-----u-- ---------- --c-----a- ----a----- ---a--a
P1/Mahoney      -a-----u-- ---------- --c-----a- ----a----- ---a--a
P2/Sabin        -a-----u-- ----c----- -----u---- -g--a----- ---a--a
P2/Lansing      -g-----c-- ---uc----c --a------- -g-----u-- -------
P2/W2           -g-----c-- ---uc----c --a------- ug-----u-- -------
P3/Leon         -u-----a-- g--------- ---------- ---------- -------
P3/Sabin        -u-----a-- g--------- ---------- ---------- -------
P3/Finland      -a--u--u-- ----c-u--- --a--u---- ---------- -------
P3/Leon 119     -u-----a-- g--------- ---------- ---------- -------
CONSENSUS       C-GACAU-AG AGACUUGUAU GCUUACGAGG AAGAGGCCAU GGAGCAG
```

APPENDIX 2 A comparison of nucleotide sequences of the region coding for the protease 2A. Complete sequences were taken from the GenBank/EMBL database and aligned using GCG 'pileup' and 'pretty' programs. Conserved residues are shown as dashes (-) and differences to the consensus are in lower case. Numbering is from type 3 Leon. The virulent revertant of Sabin 3, Leon 119 is also displayed to confirm the sequence is identical to Leon in this region.

REFERENCES

REFERNCES

Abraham, R., Minor, P., Dunn, G., Modlin, J. F. & Ogra, P. L. 1993. Shedding of virulent poliovirus revertants during immunization with oral poliovirus vaccine after prior immunization with inactivated polio vaccine. J. Infec. Dis. **168**:1105-1109.

Alexander, L., Lu, H. H. & Wimmer, E. 1994. Polioviruses containing picornavirus type 1 and/or type 2 internal ribosomal entry site elements: Genetic hybrids and the expression of a foreign gene. Proc. Natl. Acad. Sci. USA **91**:1406-1410.

Andino, R., Rieckhof, G. E. & Baltimore, D. 1990. A functional ribonucleoprotein complex forms around the 5' end of poliovirus RNA. Cell **63**:369-380.

Andino, R., Rieckhof, G. E., Achacoso, P. L. & Baltimore, D. 1993. Poliovirus RNA synthesis utilizes an RNP complex formed around the 5'-end of viral RNA. EMBO J. **12**:3587-3598.

Armstrong, C. 1939a. The experimental transmission of poliomyelitis to the Eastern Cotton Rat, *Sigmodon Hispidus Hispidus*. Pub. Health Rep. **54**:1719-1721.

Armstrong, C. 1939b. Successful transfer of the Lansing strain of poliomyelitis virus from the Cotton Rat to the White Mouse. Pub. Health Rep. **54**:2303-2305.

Attal, J., Théron, M.-C., Taboit, F., Cajero-Juarez, M., Kann, G., Bolifraud, P. & Houdebine, L. M. 1996. The RU5 ('R') region from human leukaemia viruses (HTLV-1) contains an internal ribosome entry site (IRES)-like sequence. FEBS Letts. **392**:220-224.

Auvinen, P. & Hyypiä, T. 1990. Echoviruses include genetically distinct serotypes. J. Gen. Virol. **71**:2133-2139.

Barco, A. & Carrasco, L. 1995. A human virus protein, poliovirus 2BC, induces membrane proliferation and blocks the exocytic pathway in the yeast *Saccharomyces cerevisiae*. EMBO J. **14**:3349-3364.

Barnes, W. M. 1994. PCR amplification of up to 35-kb DNA with high fidelity and high yield from λ bacteriophage templates. Proc. Natl. Acad. Sci. USA **91**:2216-2220.

Barron, A. L., Olshevsky, C. & Cohen, M. M. 1970. Characteristics of the BGM line of cells from African green monkey kidney. Arch. G. Vir. **32**:389-392.

Barton, D. J., Black, E. P. & Flanegan, J. B. 1995. Complete replication of poliovirus *in vitro*. Preinitiation RNA replication complexes require soluble cellular factors for the synthesis of VPg-linked RNA. J. Virol. **69**:5516-5527.

Bazan, J. F. & Fletterick, R. J. 1988. Viral cysteine proteases are homologous to the trypsin-like family of serine proteases: Structural and functional implications. Proc. Natl. Acad. Sci. USA **85**:7872-7876.

Belsham, G. J. & Sonenberg, N. 1996. RNA-protein interactions in regulation of picornavirus RNA translation. Microbio. Revs. **60**:499-511.

Bergelson, J. M., Cunningham, J. A., Droguett, G., Kurt-Jones, E. A., Krithivas, A., Hong, J. S., Horwitz, M. S., Crowell, R. L. & Finberg, R. W. 1997. Isolation of a common receptor for coxsackie B virus and adenoviruses 2 and 5. Science **275**:1320-1323.

Bienkowska-Szewcyzk, K. & Ehrenfeld, E. 1988. An internal 5'-noncoding region required for translation of poliovirus RNA *in vitro*. J. Virol. **62:**3068-3072.

Bienz, K., Egger, D. & Pasamontes, L. 1987. Association of polioviral proteins of the P2 genomic region with the viral replication complex and virus induced membrane synthesis as visualised by electronmicroscopy immunocytochemistry and autoradiography. Virology **160**:220-226.

Black, T. L., Safer, B., Hovanessian, A. & Katze. M. G. 1989. The cellular 68,000-M_r protein kinase is highly autophosphorylated and activated yet significantly degraded during poliovirus infection: Implications for translation regulation. J. Virol. **63**:2244-2251.

Bonneau, A.-M. & Sonenberg, N. 1987a. Proteolysis of the p220 component of the cap-binding protein complex is not sufficient for complete inhibition of host cell protein synthesis after poliovirus infection. J. Virol. **61**:986-991.

Bonneau, A.-M. & Sonenberg, N. 1987b. Involvement of the 24kDa cap-binding protein in regulation of protein synthesis in mitosis. J. Biol. Chem. **262**:11134-11139.

Borman, A., Howell, M. T., Palton, J. G. & Jackson, R. J. 1993. The involvement of a spliceosome component in internal initiation of human rhinovirus RNA translation. J. Gen. Virol. **74**:1775-1788.

Bouchard, M. J., Du-Hung, L. & Racaniello, V. 1995. Determinants of attenuation and temperature sensitivity in the type 1 poliovirus Sabin vaccine. J. Virol. **69**:4972-4978.

Brown, B. A. & Ehrenfeld, E. 1979. Translation of poliovirus RNA *in vitro*: Changes in cleavage pattern and initiation sites by ribosomal salt wash. Virology **97**:396-405.

Brown, A. B., Zhang, H., Ping, L.-H. & Lemon, S. M. 1992. Secondary structure of the 5'nontranslated regions of hepatitis C virus and pestivirus genomic RNAs. Nuc. Acids Res. **20**:5041-5045.

Brown, E. A., Zajac, A. J. & Lemon, S. M. 1994. *In vitro* characterisation of an internal ribosome entry site (IRES) present within the 5'nontranslated region of hepatitis A virus RNA: Comparison with the IRES of encephalomyocarditis virus. J. Virol. **68**:1066-1074.

Brunel, F., Zakin, M. M., Buc, H. & Buckle, M. 1996. The polypyrimidine tract binding (PTB) protein interacts with single-stranded DNA in a sequence-specific manner. Nuc. Acids Res. **24**:1608-1015.

Callahan, P. L., Mizutani, S. & Colonno, R. J. 1985. Molecular cloning and complete sequence determination of RNA genome of human rhinovirus type 14. Proc. Natl. Acad. Sci. USA **82**:732-736.

Caliguiri, L. A. & Tamm, I. 1969. Membranous structures associated with translation and transcription of poliovirus RNA. Science **166**:885-886.

Cammack, N., Phillips, A., Dunn, G., Patel, V. & Minor, P. D. 1988. Intertypic rearrangements of poliovirus strains in vaccinees. Virology **167**:507-514.

Cann, A. J., Stanway, G., Hughes, P. J., Minor, P. D., Evans, D. M. A., Schild, G. C. & Almond, J. W. 1984. Reversion to virulence of the live attenuated Sabin type 3 oral poliovirus vaccine. Nuc. Acids Res. **12**:7787-7792.

Chang, K. H., Auvinen, P., Hyypiä, T. & Stanway, G. 1989. The nucleotide sequence of coxsackievirus A9; implications for receptor binding and enterovirus classification. J. Gen Virol. **70**:3269-3280.

Chen, C-Y. & Sarnow, P. 1995. Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. Science **268**:415-417.

Chow, M., Newman, J. F. E., Filman, D., Hogle, J. M., Rowlands, D. J. & Brown, F. 1987. Myristolylation of picornavirus capsid protein VP4 and its structural significance. Nature **327**:482-486.

Christodoulou, C., Colbere-Garapin, F., Macadam, A. J., Taffs, L. F., Marsden,S., Minor, P. D. & Horaud, F. 1990. Mapping of mutations associated with neurovirulence in monkeys infected with Sabin 1 poliovirus revertants selected at high temperature. J. Virol. **64**:4922-4929.

Chumakov, K. M., Powers, L. B., Noonan, K. E., Rononseon, I. R. & Levenbrook, I. S. 1991. Correlation between amount of virus with altered sequence and the monkey test for acceptability of oral poliovirus vaccine. Proc. Natl. Acad. Sci. USA **88**:199-203.

Clarke, M. E., Hämmerle, T., Wimmer, E. & Dasgupta, A. 1991. Poliovirus proteinase 3C converts an active form of transcription factor IIIC to an inactive form: a mechanism for inhibition of host cell polymerse III transcription by poliovirus. EMBO J. **10**:2941-2947.

Coller, B.-A. G., Chapman, N. M., Beck, M. A., Pallansch, M. A., Gauntt, C. J. & Tracy, S. M. 1990. Echovirus 22 is an atypical enterovirus. J. Virol. **64**:2692-2701.

Coller, B.-A, G., Tracy, S. M. & Etchinson, D. 1991. Cap-binding complex protein p220 is not cleaved during echovirus 22 replication in HeLa cells. J. Virol. **65**:3903-3905.

Couderc, T., Christodoulou, C., Kopecka, H., Marsden, S., Taffs, L. F., Crainic, R. & Horaud. F. 1989. Molecular pathogenesis of neural lesions induced by poliovirus type 1. J. Gen. Virol. **70**:2907-2918.

Cullen, B. R. 1992. Mechanism of action of regulatory proteins encoded by complex retroviruses. Microb. Revs. **56**:375-394.

Currey, K. M., Peterlin, B. M., & Maisel, J. V. Jr. 1986. Secondary structure of poliovirus RNA: Correlation of computer-predicted with electron microscopically observed structure. Virology **148**:33-46.

Currey, S., Chow, M. & Hogle, J. M. 1996. The poliovirus 135S particle is infectious. J. Virol. **70**:7125-7131.

Dahling, D. R., Berg, G. & Berman, D. 1974. BGM, a continuous cell line more sensitive than primary rhesus and african green kidney cells for the recovery of viruses from water. Health Lab. Sci. **11**:275-282.

Davies, M. V., Pelletier, J., Meerovitch, K., Sonenberg, N. & Kaufman, R. J. The effect of of poliovirus proteinase $2A^{PRO}$ expression on cellular metabolism. 1991. J. Biol. Chem. **266**:14714-14720.

Devaney, M. A., Vakharia, V. N., Lloyd, R. E., Ehrenfeld, E. & Grubmen, M. J. 1988. Leader protein of foot-and-mouth disease virus is required for cleavage of the p220 component of the cap-binding protein complex. J. Virol. **62**:4407-4409.

Dorner, A. J., Semler, B. L., Jackson, R. J., Hanecak, R., Duprey, E. & Wimmer, E. 1984. *In vitro* translation of poliovirus RNA: Utilization of internal initiation sites in reticulocyte lysate. J. Virol. **50**:507-514.

Duechler, M., Skern, T., Sommergruber, W., Neubauer, C., Gruendler, P., Fogy, I., Blaas, D. & Kuechler, E. 1987. Evolutionary relationships within the human rhinovirus genus: Comparison of serotypes 89, 2 and 14. Proc. Natl. Acad. Sci. USA **84**:2605-2609.

Duke, G. M., Hoffman, M. A. & Palmenberg, A. C. 1992. Sequence and structural elements that contribute to efficient encephalomyocarditis virus RNA translation. J. Virol. **66**:1602-1609.

Earle, J. A. P., Skuce, R. A., Flemming, C. S., Hoey, E. M. & Martin, S. J. 1988. The complete nucleotide sequence of a Bovine Enterovirus. J. Gen. Virol. **69**:253-263.

Electroporation system operation manual. 1991. Published by Biotechnologies & Experimental Research Inc.

Etchison, D., Milburn, S. C., Edery, I., Sonenberg, N. & Hershey, J. B. W. 1982. Inhibition of HeLa cell protein synthesis following poliovirus infection correlates with the proteolysis of a 220,000-dalton polypeptide associated with eucaryotic initiation factor 3 and a cap binding protein complex. J. Biol. Chem. **257**:14806-14810.

Etchison, D. & Fout, S. 1985. Human rhinovirus 14 infection of HeLa cells results in the proteolytic cleavage of the p220 cap-binding complex subunit and inactivates Globin mRNA translation *in vitro*. J. Virol. **54**:634-638.

Evans, D. M A., Dunn, G., Minor, P. D., Schild, G. C., Cann, A. J., Stanway, G., Almond, G. W., Currey, K. & Maizel, J. V. 1985. A single nucleotide change in the 5' non coding region of the genome of the Sabin type 3 poliovaccine is associated with increased neurovirulence. Nature **314**:548-550.

Flanegan J. B. & Van Dyke, T. A. 1979. Isolation of a soluble and template-dependent poliovirus RNA polymerase that copies virion RNA *in vitro*. J. Virol. **32**:155-161.

Francki, R. I. B., Fauquet, C. M., Knudson, D. L. & Brown, F. 1994. Classification et nomenclature des virus. Pub. Springer-Verlag France.

Fricks, C. E. & Hogle, J. M. 1990. Cell-induced conformational change in poliovirus: Externalization of the amino terminus of VP1 is responsible for liposome binding. J. Virol. **64**:1934-1945.

Fukushi, S., Katayama, K., Kurihara, C., Ishiyama, N., Hoshino, F. B., Ando, T. & Oya, A. 1994. Complete 5' noncoding region is necessary for the efficient internal initiation of hepatitis C virus RNA. Biochem. Biophys. Res. Commun. **199**:425-432.

Gan, W. & Rhoads, R. E. 1996. Internal initiation of translation directed by the 5'-untranslated region of the mRNA for eIF4G, a factor involved in the picornavirus-induced switch from cap-dependent translation to internal initiation. J. Biol. Chem. **271**:623-626.

Genetics Computer Group. 1994. Program manual for the Wisconsin PAckage, Version 8, Genetics Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711.

Giachetti, C. & Semler, B. L. 1991. Role of viral membrane polypeptide in strand-specific initiation of poliovirus RNA synthesis. J. Virol. **65**:2647-2654.

Gingras, A.-C, Svitkin, Y., Belsham, G. J., Pause, A. & Sonenberg, N. 1996. Activation of the translational suppressor 4E-BP1 following infection with encephalomyocarditis virus and poliovirus. Proc. Natl. Acad. Sci. USA **93**:5578-5583.

Gottleib, E. & Steitz, J. A. 1989. The RNA binding protein La influences both the accuracy and the efficiency of RNA polymerase III transcription *in vitro*. EMBO J. **8**:841-850.

Greve, J. M., Davis, G., Marlor, A. M., Forte, C. P., Yost, S. C., Marlor, C. W., Kamarch, M. E. & McClelland, A. 1989. The major human rhinovirus receptor is ICAM-1. Cell **56**:839-847.

Guinea, R. & Carrasco, L. 1990. Phospholipid biosynthesis and poliovirus genome replication, two coupled phenomena. EMBO J. **9**:2011-2016.

Guttman, N. & Baltimore, D. 1977. A plasma membrane component able to bind and alter virions of poliovirus type 1: Studies on cell-free alteration using a simplified assay. Virology **82**:25-36.

Haller, A. A. &Semler, B. L. 1992. Linker scanning mutagenesis of the internal ribosome entry site of poliovirus RNA. J, Virol. **66**:5075-5086.

Haller, A. A., Nguyen, J. H. C. & Semler, B. L. 1993. Minimum internal ribosome entry site required for poliovirus infectivity. J. Virol. **67**:7461-7471.

Haller, A. A., Stewart, S. R. & Semler, B. L. 1996. Attenuation stem-loop lesions in the 5' noncoding region of poliovirus RNA: Neuronal cell-specific translation defects. J. Virol. **70**:1467-1474.

Hambidge, S. J. & Sarnow, P. 1992. Translational enhancement of the poliovirus 5'noncoding region mediated by virus-encoded polypeptide 2A. Proc. Natl. Acad. Sci. USA **89**:10272-10276.

Hashimoto, I., Hagiwara, A. & Komatsu, T. 1984. Ultrastructural studies on the pathogenesis of poliomyelitis in monkeys infected with poliovirus. Acta Neuropathol. (Berlin) **64**:53-60.

Hellen, C. U. T., Fäcke, M., Kräusslich, H., Lee, C. & Wimmer, E. 1991. Characterisaiton of poliovirus 2A proteinase by mutational analysis: residues required for autocatalytic activity are essential for induction of cleavage by eukaryotic initiation factor polypeptide p220. J. Virol. **65**:4226-4231.

Hellen, C. U. T., Witherell, G. W., Schmid, M., Hoon Shin, S., Pestova, T. V., Gil, A. & Wimmer, E. 1993. A cytoplasmic 57kDa protein that is required for translation of picornavirus RNA by internal ribosome entry is identical to the nuclear pyrimidine tract-binding protein. Proc. Natl. Acad. Sci. USA **90**:7642-7646.

Hellen, C. U. T., Pestova, T. V., Litterst, M. & Wimmer, E. 1994. The cellular polypeptide p57 (pyrimidine tract-binding protein) binds to multiple sites in the poliovirus 5'nontranslated region. J. Virol. **68**:941-950.

Hewlett, M. J., Rose, J. K. & Baltimore, D. 1976. 5' terminal structure of poliovirus polyribosomal RNA is pUp. Proc. Natl. Acad. Sci. USA **73**:327-330.

Hogle, J. M., Chow, M. & Filman, D. J. 1985. Three dimensional structure of poliovirus at 2.9Å resolution. Sci. **229**:1358-1365.

Holland, J. J. 1962. Irreversible eclipse of poliovirus by HeLa cells. Virology **16**:163-176.

Honda, M., Ping, L.-H., Runbrand, R. C. A., Amphlett, E., Clarke, B., Rowland, D. J. & Lemon, S. M. 1996. Structural requirements for initiation of translation by internal ribosome entry within genome-length hepatitis C virus RNA. Virology **222**:31-42.

Huang, J. & Schneider, R. J. 1991. Adenovirus inhibition of cellular protein synthesis involves inactivation of cap-binding protein. Cell **65**:271-280.

Hughes, P. J., Evans, D. M. A., Minor, P. D., Schild, G. C., Almond, J. W. & Stanway, G. 1986. The nucleotide sequence of a type 3 poliovirus isolated during a recent outbreak of poliomyelitis in Finland. J. Gen. Virol. **67**:2093-2102.

Hughes, P. J., North, C., Jellis, C. H., Minor, P. D. & Stanway, G. 1988. The nucleotide sequence of human rhinovirus 1B: molecular relationships within the rhinovirus genus. J. Gen. Virol. **69**:49-58.

Hughes, P. J., North, C., Minor, P. D. & Stanway, G. 1989. The complete nucleotide sequence of coxsackievirus A21. J. Gen. Virol. **70**:2943-2952.

Iizuker, N., Kuge, S. & Nomoto, A. 1987. Complete nucleotide sequence of the genome of coxsackievirus B1. Virology. **156**:64-73.

Iizuka, N., Kohara, M., Hagino-Yamagishi, K., Abe, S., Komatsu, T., Tago, K., Arita, M. & Nomoto, A. 1989. Construction of less neurovirulent polioviruses by introducing deletions into the 5' noncoding sequence of the genome. J. Virol. **63**:5354-5363.

Iizuka, N., Najita, L., Franzusoff, A. & Sarnow, P. 1994. Cap-dependent and cap-independent translation by internal initiation of mRNAs in cell extracts prepared from *Saccaromyces cerevisiae*. Mol. Cell. Biol. **14**:7322-7330.

Inoue, T., Suzuki, T. & Sekiguchi, K. 1989. The complete nucleotide sequence of swine vesicular disease virus. J. Gen. Virol. **70**:919-934.

Ishitsuka, H., Ninomiya, Y. T., Ohsawa, C., Fujiu, M. & Suhara, Y. 1982a. Direct and specific inactivation of rhinovirus by chalcone Ro 09-0410. Antimicrob. Agents Chemother. **22**:617-621.

Ishitsuka, H., Ohsawa, C., Ohiwa, T., Umeda, I. & Suhara, T. 1982b. Antipicornavirus flavone Ro 09-0179. Antimicrob. Agents Chemother. **22**:611-616.

Jackson, R. J., Hunt, S. L., Gibbs, C. L. & Kaminski, A. 1994. Internal initiation of translation of picornavirus RNAs. Mol. Biol. Reps. **19**:147-159.

Jacobs, J. P., Jones, C. M. & Baille, J. P. 1970. Characteristics of a human diploid cell designated MRC-5. Nature **227**:168-170.

Jang, S. K., Kräusslich, H.-G., Nicklin, M. J. H., Duke, G. M., Palmenberg, A. C. & Wimmer, E. 1988. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during *in vitro* translation. J. Virol. **62**:2636-2643.

Jenkins, O., Booth, J. D., Minor, P. D. & Almond, J. W. 1987. The complete nucleotide sequence of coxsackievirus B4 and its comparison to other members of the picornaviridae. J. Gen. Virol. 1987. **68**:1835-1848.

Johnson, K. L. & Sarnow, P. 1991. Three poliovirus 2B mutants exhibit noncomplementable defects in viral RNA amplification and display dosage-dependent dominance over wild type poliovirus. J. Virol. **65**:4341-4349.

Jore, J., de Geus, B., Jackson, R. J., Pouwles, P. H. & Enger-valk, B. E. 1988. Poliovirus protein 3CD is the active protease for processing of the precusor protein P1 *in vitro*. J. Gen. Virol. **69**:1627-1636.

Jubelt, B., Ropka, S. L., Goldfarb, S., Waltenbaugh, C. & Oates, R. P. 1991. Susceptability and resistance to poliovirus-induced paralysis of inbred mouse strains. J. Virol. **65**:1035-1040.

Kaminski, A., Howell, M. T. & Jackson, R. J. 1990. Initiation of encephalomyocarditis virus RNA translation: the authentic initiation site is not selected by a scanning mechanism. EMBO J. **9**:3753-3759.

Kawamura, N., Kohara, M., Abe, S., Komatsu, T., Tago, K., Arita, M. & Nomoto, A. 1989. Determinants in the 5'noncoding region of poliovirus Sabin 1 RNA that influence the attenuation phenotype. J. Virol. **63**:1302-1309.

Kenan, D. J., Query, C. C. & Keene, J. D. 1991. RNA recognition: Towards identifying determinants of specificity. Trends in Biochem. Sci. **16**:214-220.

Kirchweger, R., Ziegler, E., Lamphear, B. J., Wates, D., Liebig, H.-D., Sommerguber, W., Sobrino, F., Hohenadl, C., Blaas, D., Rhoads, R. E. & Skern, T. 1994. Foot-and-mouth disease virus leader proteinase: Purification of the Lb form and determination of its cleavage site on eIF-4γ. J. Virol. **68**:5677-5684.

Kitamura, M., Semler, B. L., Rothberg, P. G., Larsen, G. R., Adler, C. J., Doner, A. J., Emini, E. A., Henecak, R., Lee, J. J., van der Werf, S., Anderson, C. W. & Wimmer, E. 1981. Primary structure, gene organisation and polypeptide expression of poliovirus RNA. Nature **291**:482-485.

Klump, W. M., Bergmann, I., Müller, B. C., Ameis, D. & Kandolf, R. 1990. Complete nucleotide sequence of infectious coxsackievirus B3 cDNA: two initial 5' uridine residues are regained during plus-strand RNA synthesis. J. Virol. **64**:1573-1583.

Koike, S., Ise, I., Sato. Y., Yonekawa, H., Gotoh, O. & Nomoto, A. 1992. A second gene for the African green monkey poliovirus receptor that has no putative N-glycosylation site in the functional N-terminal immunoglobulin-like domain. J. Virol. **66**:7059-7066.

Kozak, M. 1978. How do eucaryotic ribosomes select inititation regions in messenger RNA? Cell **15**:1109-1123.

Kozak, M. 1989. The scanning model for translation: An update. J. Cell Biol. **108**:229-241.

Kraus, W., Zimmermann, H., Zimmermann, A., Eggers, H. J. & Nelsen-Salz, B. 1995. Infectious cDNA clones of echovirus 12 and a variant resistant against the uncoating inhibitor rhodanine differ in seven amino acids. J. Virol. **69**:5853-5858.

Kräusslich, H. G., Nicklin, M. J. H., Toyoda, H., Etchison, D. & Wimmer, E. 1987. Poliovirus proteinase 2A induces cleavage of eucaryotic initiation factor 4F polypeptide p220. J. Virol. **61**:2111-2718.

Kräusslich, H. G., Hölscher, C., Reuer, Q., Harber, J. & Wimmer, E. 1990. Myristoylation of the poliovirus polyprotein is required for proteolytic processing of the capsid and for viral infectivity. J. Virol. **64**:2433-2436.

Kuge, S., Kawamura, N. & Nomoto, A. 1989. Strong inclination toward transition mutation in nucleotide substitutions by poliovirus replicase. J. Mol. Biol. **207**:175-182.

Lama, J., Sanz, M. A. & Rodríguez, P. L. 1995. A role for 2AB protein in poliovirus genome replication. J. Biol. Chem. **270**:14430-14438.

Lamphear, B. J., Yan, R., Yang, F., Aters, D., Liebig, H.-D., Klump, H., Kuechler, E., Skern, T. & Rhoads, R. E. 1993. Mapping the cleavage site in protein synthesis initiation factor eIF-4γ of the 2A proteases from human coxackievirus and rhinovirus. J. Biol. Chem. **268**:19200-19203.

Lamphear, B. J., Kirchweger, R., Skern, T. & Rhoads, R. E. 1995. Mapping of functional domains in eukaryotic protein synthesis initiation factor 4G (eIF4G) with picornaviral proteases. J. Biol. Chem. **270**:21975-21983.

La Monica, N., Meriam, C. & Racaniello, R. 1986. Mapping of sequences required for mouse neurovirulence of poliovirus type 2 Lansing. J. Virol. **57**:515-525.

Le, S.-Y. & Zuker, M. 1990. Common structures of the 5' non-coding region in enteroviruses and rhinoviruses. Thermodynamical stability and statistical significance. J. Mol. Biol. **216**:729-741.

Le, S.-Y., Chen, J.-H., Sonenberg, N. & Maizel, J. V., Jr. 1992. Conserved tertiary structure elements in the 5' untranslated region of human enteroviruses and rhinoviruses. Virology **191**:858-866.

Le, S.-Y., Chen, J-H., Sonenberg, N. & Maizel, J. V., Jr. 1993. Conserved tertiary structural elements in the 5' nontranslated region of cardiovirus, aphthovirus and hepatitis A virus RNAs. Nuc. Acids Res. **21**:2445-2451.

Lee, S.-Y., & Maizel, J. V. Jr. 1997. A common RNA structural motif involved in the internal initiation of translation of cellular mRNAs. Nuc. Acids Res. **25**:362-369.

Lee, C. & Wimmer, E. 1988. Proteolytic processing of poliovirus polyprotein: elimination of $2A^{PRO}$ mediated alternative cleavage of polypeptide 3CD by *in vitro* mutagensis. Virology **166**:405-414.

Lee, W.-M., Wang, W. & Rueckert, R. R. 1994. Complete sequence of the RNA genome of human rhinovirus 16, a clinically useful common cold virus belonging to the ICAM-1 receptor group. Virus Genes **9**:177-181.

Li, J.-P. & Baltimore, D. 1988. Isolation of poliovirus 2C mutants defective in viral RNA synthesis. J. Virol. **62**:4016-4021.

Li, J.-P. & Baltimore, D. 1990. An intragenic revertant of a poliovirus 2C mutant has an uncoating defect. J. Virol. **64**:1102-1107.

Liebig, H.-D., Ziegler, E., Yan, R., Hartmuth, K., Klump, H., Kowalski, H., Blaas, D., Sommergruber, W., Frasel, L., Lamphear, B., Rhoads, R., Kuechler, E. & Skern, T. 1993. Purification of two picornaviral 2A proteinases: Interaction with eIf-4γ and influence on *in vitro* translation. **32**:7581-7585.

Lindberg, A. M., Stålhandske, P. O. K. & Pettersson, U. 1987. Genome of coxsackievirus B3. Virology **156**:50-63.

Liu, D. X. & Inglis, S. S. 1992. Internal entry of ribosomes on a tricistronic mRNA encoded by infectious bronchitis virus. J. Virol. **66**:6143-6154.

Lloyd, R. E., Toyoda, H., Etchison, D., Wimmer, E. & Ehrenfeld, E. 1986. Cleavage of the cap binding protein complex polypeptide p220 is not effected by the second poliovirus protease 2A. Virology **150**:299-303.

Lu, H.-H., Yang, C.-F., Murdin, A. D., Klein, M. H., Harber, J. J., Kew, O. M. & Wimmer, E. 1994. Mouse neurovirulence determinants of poliovirus type 1 strain LS-a map to the coding regions of capsid protein VP1 and proteinase $2A^{PRO}$. J. Virol. **68**:7507-7515.

Luz, N. & Beck, E. 1991. Interaction of cellular 57-kilodalton protein with the internal translation initiation site of foot-and-mouth disease virus. J. Virol. **65**:6486-6496.

Macadam, A. J., Arnold, C., Howlett, J., John, A., Marsden, S., Taffs, F., Reeve, P., Hamada, N., Wareham, K., Almond, J., Cammack, N. & Minor, P. D. 1989. Reversion of the attenuated and temperature-sensitive phenotypes of the Sabin type 3 strain of poliovirus in vaccinees. Virology **172**:408-414.

Macadam, A. J., Ferguson, G., Arnold, C. & Minor, P. D. 1991a. An assembly defect as a result of an attenuating mutation in the capsid proteins of the poliovirus type 3 vaccine strain. J. Virol. **65**:5225-5231.

Macadam, A. J., Pollard, S. R., Ferguson, G., Dunn, G., Skuce, R., Almond, J. W. & Minor, P. D. 1991b. The 5'noncoding region of the type 2 poliovirus vaccine strain contains determinants of attenuation and temperature sensitivity. Virology. **181**:451-458.

Macadam, A. J., Ferguson, G., Burlison, J., Stone, D., Skuce, R., Almond, J. W. & Minor, P. D. 1992. Correlation of RNA secondary structure and attenuation of Sabin vaccine strains of poliovirus in tissue culture. Virology **189**:415-422.

Macadam, A. J., Pollard, S. R., Ferguson, G., Skuce, R., Wood, D., Almond, J. W. & Minor, P. D. 1993. Genetic basis of attenuation of the Sabin type 2 vaccine strain of poliovirus in primates. Virology **192**:18-26.

Macadam, A. J., Ferguson, G., Fleming, T., Stone, D. M., Almond, J. W. & Minor, P. M. 1994. Role for poliovirus protease 2A in cap-independent translation. EMBO J. **13:** 924-927.

Macejak, D. G. & Sarnow, P. 1991. Internal initiation of translation mediated by the 5' leader of a cellular mRNA. Nature **353**:90-94.

Maddon, P. J., Dalgleish, A. G., McDougal, J. S., Clapham, P. R., Weiss, R. A. & Axel, R. 1986. The T4 gene encodes the AIDS virus receptor and is expressed in the immune system and the brain. Cell **47**:333-348.

Marc, D., Drugeon, G., Haenne, A-L., Girard, M. & van der Werf, S. 1989. Role of myristoylation of poliovirus capsid protein VP4 as determined by site-directed mutagenesis of its N-terminal sequence. EMBO J. **8**:2661-2668.

Mason, P. W., Piccone, M. E., McKenna, St.-C., Chinsangaram, J. & Grubman, M. J. 1997. Evaluation of a live-attenuated foot-and-mouth disease virus as a vaccine candidate. Virology **227**:96-102.

McGoldrick, A., Macadam, A. J., Dunn, G., Rowe, A., Burlison, J., Minor, P. D., Meredith, J., Evans, D. J. & Almond, J. W. 1995. Role of mutations G-480 and C-6203 in the attenuation phenotype of Sabin type 1 poliovirus. J. Virol. **69**:7601-7605.

Medina, M., Domingo, E., Brangwyn, J. K. & Belsham, G. J. 1993. The two species of the foot-and-mouth disease virus leader protein, expressed individually, exhibit the same activities. Virology **194**:355-359.

Mendelsohn, C., Johnson, B., Lionetti, K. A., Nobis, P., Wimmer, E. & Raciello, V. R. 1986. Transformation of a human poliovirus receptor gene into mouse cells. Proc. Natl. Acad. Sci. USA **83**:7845-7849.

Mendelsohn, C. L., Wimmer, E. & Racaniello, V. R. 1989. Cellular receptor for poliovirus: Molecular cloning, nucleotide sequence and expression of a new member of the immunoglobulin superfamily. Cell **56**:855-865.

Meerovitch, K., Pelletier, J. & Sonenberg, N. 1989. A cellular protein that binds to the 5'-noncoding region of poliovirus RNA: implications for internal translation initiation. Genes Dev. **3**:1026-1034.

Meerovitch, K., Svitkin, Y. V., Lee, H. S., Lejbkowicz, F., Kenan, D. J., Chan, E. K. L., Agol, V. I., Keene, J. D. & Sonenberg, N. 1993. La autoantigen enhances and corrects aberrant translation of poliovirus RNA in reticulocyte lysate. J. Virol. **67**:3798-3807.

Minor, P. D. 1980. Comparative biochemical studies of type 3 poliovirus. J. Virol. **34**:73-84.

Minor, P. D., John, A., Ferguson, M. & Icenogle, J. P. 1986. Antigenic and molecular evolution of the vaccine strain of type 3 poliovirus during the period of excretion by a primary vaccinee. J. Gen. Virol. **67**:693-706.

Minor, P. D. & Dunn, G. 1988. Effect of sequences in the 5' non-coding region on the replication of polioviruses in the human gut. J. Virol. **69**:1091-1096.

Minor, P. D. 1992. The molecular biology of poliovaccines. J. Gen. Virol. **73**:3065-3077.

Molla, A., Hellen, C. U. T. & Wimmer, E. 1993a. Inhibition of proteolytic activity of poliovirus and rhinovirus 2A proteinases by elastase-specific inhibitors. J. Virol. **67**:4688-4695.

Molla, A., Paul, A. V., Schmid, M., Jang, S. K. & Wimmer, E. 1993b. Studies on dicistronic polioviruses implicate viral proteinase $2A^{PRO}$ in RNA replication. Virology **196**:739-747.

Moore, A. E., Sabachewsky, L. & Toolan, H. W. 1955. Culture characteristics of four permanent lines of human cancer cells. Cancer Res. **15**:598-605.

Mosenkis, J., Daniels-McQueen, S., Janovec, S., Duncan, R., Hershey, J. W. B., Grifo, J. A., Merrick, W. C. & Thack, R. E. 1985. Shut off of host translation by encephalomyocarditis virus infection does not involve cleavage of the eucaryotic initiation factor 4F polypeptide that accompanies poliovirus infection. J. Virol. **54**:643-645.

Mosser, A. G., Caliguiri, L. A. & Tamm, I. 1972. Incorporation of lipid precursors into cytoplasmic membranes of poliovirus infected HeLa cells. Virology **47**:39-47.

Muzychenko, A. R., Lipskaya, G. Y., Maslova, S. V., Svitkin, Y. V., Pilipenko, E. V., Nottay, B. K., Kew, O. M. & Agol, V. I. 1991. Coupled mutations in the 5'-untranslated region of the Sabin poliovirus strains during *in vivo* passages: structural and functional implications. Virus Res. **21**:111-122.

Nathanson, N. & Langmuir, A. D. 1963. The cutter incident. Poliomyelitis following formaldehyde-inactivated poliovirus vaccination in the United States during the spring of 1955. Am. J. Hyg. **78**:16-28.

Nicholson, R., Pelletier, J., Le, S-Y. Sonenberg, N. 1991. Structural and functional analysis of the ribosome landing pad of poliovirus type 2: In Vivo translation studies. J. Virol. **65**:5886-5894.

Nkowane, B. M., Wassilak, S. G. F., Oversteen, W. A., Bart, K. J., Schonberger, L. B., Hinman, A. R. & Kew, O. M. 1987. Vaccine-associated paralytic poliomyelitis in the United States: 1973 through 1984. J. Am. Med. Assoc. **257**:1335-1340.

Nomoto, A., Lee, Y. F. & Wimmer, E. 1976. The 5' end of the poliovirus mRNA is not capped with $m^7G(5')Np$. Proc. Natl. Acad. Sci. USA **73:**375-380.

Nomoto, A., Omata, T., Toyoda, H., Kuge, S., Horie, H., Kataoka, Y., Genba, Y., Nakano, Y. & Imura, N. 1982. Complete nucleotide sequence of the attenuated poliovirus Sabin 1 strain genome. Proc. Natl. Acad. Sci. USA **79**:5793-5787.

Oh, S. K., Scott, M. P. & Sarnow, P. 1992. Homeotic gene antennapedia messenger-RNA contains 5'non-coding sequences that confer translational initiation by internal ribosome binding. Genes Dev. **6**:1643-1653.

Ohlmann, T., Rau, M., Morley, S. J. & Pain, V. M. 1995. Proteolytic cleavage of initiation factor eIF-4γ in the reticulocyte lysate inhibits translation of capped mRNAs but enhances that of uncapped mRNAs. Nuc. Acids Res. 23:334-340.

Ohlmann, T., Rau, M., Pain, V. M. & Morley, S. J. 1996. The C-terminal domain of eukaryotic protein synthesis initiation factor (eIF) 4G is sufficient to support cap-independent translation in the absence of eIF4E. EMBO J. **15**:1371-1382.

Olsen, N. H., Kolatkar, P. R., Oliveira, M. A., Cheng, R. H., Greve, J. M., McClelland, A., Baker, T. S. & Rossmann, M. G. 1993. Structure of a human rhinovirus complexed with its receptor molecule. Proc. Natl. Acad. Sci. USA **90**:507-511.

Omata, T., Kohara, M., Kuge, S., Komatsu, T., Abe, S., Semler, B. L., Kameda, A., Itoch, H., Arita, M., Wimmer, E. & Nomoto, A. 1986. Genetic analysis of the attenuation phenotype of poliovirus type 1. J. Virol. **58**:348-358.

O'Neill. R. E. & Racaniello, V. R. 1989. Inhibition of translation in cells infected with a poliovirus $2A^{PRO}$ mutant correlates with phosphorylation of the alpha subunit of eucaryotic initiation factor 2. J. Virol. **63**:5069-5075.

Palmenberg, A. C. 1982. *In vitro* synthesis and assembly of picornaviral capsid intermediate structures. J. Virol. **44**:900-906.

Palmenberg, A. C. 1989. Sequence alignments of picornaviral capsid proteins. In *Molecular Aspects of Picornavirus Infection and Detection*, pp 211-241. Edited by B. L. Semler & E. Ehrenfeld. Washington DC.: American Society for Microbiology.

Paul, A. V., Schultz, A., Pincus, S. E., Ororzlan, S. & Wimmer, E. 1987. Capsid protein VP4 of poliovirus is N-myristoylated. Proc. Natl. Acad. Sci. USA **84**:7827-7831.

Pause, A., Méthot, N., Svitkin, Y., Merrick, W. C. & Sonenberg, N. 1994a. Dominant negative mutants of mammalian translation initiation factor eIF-4A define a critical role for eIF-4F in cap-dependent and cap-independent initiation of translation. EMBO J. **13**:1205-1215.

Pause, A., Belsham, G. J., Gingras, A.-C., Donzé, O., Lin, T.-A., Lawrence, J. E. & Sonenberg, N. 1994b. Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5'-cap function. Nature **371**:762-767.

Pelletier, J., Kaplan, G., Racaniello, V. R. & Sonenberg, N. 1988. Translational efficiency of poliovirus mRNA: Mapping inhibitory *cis*-acting elements within the 5'noncoding region. J. Virol. **62**:2219-2227.

Pelletier, J. & Sonenberg, N. 1988. Internal intiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature **334**:320-325.

Percy, N., Belsham, G. J., Brangwyn, J. K., Sullivan, M., Stone, D. M. & Almond, J. W. 1992. Intracellular modifications induced by poliovirus reduce the requirement for structural motifs in the 5'noncoding region of the genome involved in internal initiation of protein synthesis. J. Virol. **66**:1695-1701.

Pérez, L. & Carrasco, L. 1992. Lack of direct correlation between p220 cleavage and the shut off of host translation after poliovirus infection. Virology **189**:178-186.

Pestova, T. V., Hellen, C. U. T. & Wimmer, E. 1991. Translation of poliovirus RNA: Role of an essential *cis*-acting oligopyrimidine element within the 5' nontranslated region and involvement of a cellular 57 kilodalton protein. J. Virol. **65**:6194-6204.

Pestova, T. V., Hellen, C. U. T. & Shatsky, I. N. 1996a. Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. Mol. Cell. Biol. **16**:6859-6869.

Pestova, T. V., Shatsky, I. N., & Hellen, C. U. T. 1996b. Functional dissection of eukaryotic intiation factor 4F: the 4A subunit and the central domain of the 4G subunit are sufficient to mediate internal entry of 43S preinitiation complexes. Mol. Cell. Biol. **16**:6870-6878.

Pevear, D. C., Oh, C. K., Cunningham, L. L., Calenoff, M. & Jubelt, B. 1990. Localisation of genomic regions specific for the attenuated mouse-adapted poliovirus type 2 strain w-2. J. Gen. Virol. **71**:43-52.

Phillips, B. A., & Emmert, A. 1986. Modulation of the expression of poliovirus proteins in reticulocyte lysates. Virology **148**:255-267.

Piccone, M. E., Rieder, E., Mason, P. W. & Gribman, M. J. 1995. The foot-and-mouth disease virus leader proteinase gene is not required for viral replication. J. Virol. **69**:5376-5382.

Pilipenko, E. V., Blinov, V. M., Romanova, L. I., Sinyakov, A. N., Maslova, S. V. & Agol, V. I. 1989. Conserved structural domains in the 5'-untranslated region of picornaviral genomes: An analysis of the segment controlling translation and neurovirulence. Virology **168**:201-209.

Pipkin, P. A., Wood, D. J., Racaniello, V. R. & Minor, P. D. 1993. Characterisation of L cells expressing the human poliovirus receptor for the specific detection of polioviruses in vitro. J. Virol. Met. **41**:333-340.

Pollard, S. R., Dunn, G., Cammack, N., Minor, P. D. & Almond, J. W. 1989. Nucleotide sequence of a neurovirulent variant of the type 2 oral poliovirus vaccine. J. Virol. **63**:4949-4951.

Poole, T. L., Wang, C., Popp, R. A., Potgieter, L. N. D., Siddiqui, A. & Colletti, M. S. 1995. Pestivirus translation initiation occurs by internal ribosome entry. Virology **206**:750-754.

Pöyry, T., Kinnunen, L. & Hovi, T. 1992. Genetic variation *in vivo* and proposed functional domains of the 5' noncoding region of poliovirus RNA. J. Virol. **66**:5313-5319.

Pöyry, T., Hyypiä, T., Horsnell, C., Kinnunen, L., Hovi, T. & Stanway, G. 1994. Molecular analysis of coxsackievirus A16 reveals a new genetic group of enteroviruses. Virology **202**:982-987.

Pöyry, T., Kinnunen, L., Hyypiä, T., Brown, B., Horsnell, C., Hovi, T. & Stanway, G. 1996. Genetic and phylogenetic clustering of enteroviruses. J. Gen. Virol. **77**:1699-1717.

Racaniello, V. R. & Baltimore, D. 1981a. Molecular cloning of poliovirus cDNA and determination of the complete nucleotide sequence of the viral genome. Proc. Natl. Acad. Sci. USA **78**:4887-4891.

Racaniello, V. R. & Baltimore, D. 1981b. Cloned poliovirus complementary DNA is infectious in mammalian cells. Science **214**:916-918.

Racaniello, V. R. 1984. Poliovirus type 2 produced from cloned cDNA is infectious in mice. Virus Res. **1**:669-675.

Ray, B. K., Lawson, T. G., Kramer, J. C., Cladaras, H. H., Grifo, J. A., Abramson, R. D., Merrick, W. C. & Thach, R. E. 1985. ATP-dependent unwinding of messenger RNA structure by eukaryotic initiation factors. J. Biol. Chem. **260**:7651-7658.

Ren, R., Costantini, F., Gorgacz, E. J., Lee, J. J. & Racaniello, V. R. 1990. Transgenic mice expressing a human poliovirus receptor: A new model for poliomyelitis. Cell **63**:353-362.

Ren, R., Moss, E. G. & Racaniello, V. R. 1991. Identification of two determinants that attenuate vaccine-related type 2 poliovirus. J. Virol. **65**:1377-1382.

Reynolds, J. E., Kaminski, A., Kettinen, H. J., Grace, K., Clarke, B. E., Carroll, A. R., Rowlands, D. J. & Jackson, R. J. 1995. Unique features of internal initiation of hepatitis C virus RNA translation. EMBO J. **14**:6010-6020.

Reynolds, J. E., Kaminski, A., Carroll, A. R., Clarke, B. E., Rowlands, D. J. & Jackson, R. J. 1996. Internal initiation of translation of hepatitis C virus RNA: The ribosome entry site is at the authentic intitiation codon. RNA **2**:867-878.

Rico-Hesse, R., Pallansch, M. A., Nottay, B. K. & Kew, O. M. 1987. Geographical disrtibution of wild poliovirus type 1 genotypes. Virology **160**:311-322.

Rijnbrand, R., Bredenbeek, P., van der Straaten, T., Whetter, L., Inchauspé, G., Lemon, S. & Spaan, W. 1995. Almost the entire 5' non-translated region of hepatitis C virus is required for cap independent translation. FEBS Lett. **365**:115-119.

Rivera, V. M., Welsh, J, D. & Maizel, J. V., Jr. 1988. Comparative sequence analysis of the 5'noncoding region of the enteroviruses and rhinoviruses. Virology **165**:42-50.

Roberts, L. O. & Belsham, G. J. 1997. Complementation of defective picornavirus internal ribosome entry site (IRES) elements by the coexpression of fragments of the IRES. Virology **227**:53-62.

Roehl, H. H., Parsley, T. B., Ho, T. V. & Semler, B. L. 1997. Processing of a cellular polypeptide by 3CD proteinase is required for poliovirus ribonucleoprotein complex formation. J. Virol. **71**:578-585.

Rohll, J. B., Percy, N., Ley, R., Evans, D. J., Almond, J. W. & Barclay, W. 1994. The 5'-untranslated regions of picornavirus RNAs contain independent functional domains essential for RNA replication and translation. J. Virol. **68**:4384-4391.

Rossmann, M. G., Arnold, E., Erichson, J. W., Frankenberger, E. A., Griffith, J. P., Hect, H-J., Johnson, J. E., Kramer, G., Luo, M., Mosser, A. G., Rueckert, R. R., Sherry, B. & Vriend, G. 1985. Structure of a human common cold virus and functional relationship to other picornaviruses. Nature **317**:145-153.

Rossmann, M. G. & Palmenberg, A. C. 1988. Conservation of the putative receptor attachment site in Picornaviruses. Virology **164**:373-382.

Ryan, M. D., Jenkins, O., Hughes, P. J., Brown, A., Knowles, N. J., Booth, D., Minor, P. D. & Almond, J. W. 1990. The complete nucleotide sequence of enterovirus type 70: relationships with other members of the picornaviridae. J Gen. Virol. **71**:2291-2299.

Sabin, A. B. & Boulger, L. R. 1973. History of Sabin attenuated poliovirus oral live vaccine strains. J. Biol. Stands. **1**:115-118.

Saiki, R. K., Scharf, S., Faloona, F., Horn, G. T., Erlich, H. A. & Arnheim, N. 1985. Enzymatic amplification of β-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. Science **230**:1350-1354.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**:487-491.

Sambrook, J., Fritsch, E. F. & Maniatis, T. 1989. Molecular Cloning, A Laboratory Manual, second edition. Cold Spring Harbor Laboratory Press.

Sandoval, I. V. & Carrasco, L. 1997. Poliovirus infection and expression of the poliovirus protein 2B provoke the disassembly of the golgi complex, the organelle target for the antipoliovirus drug Ro-090179. J. Virol. **71**:4679-4693.

Sanger, F., Nicklen, S. & Coulson, A. R. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463-5467.

Schultz, D. E., Hardin, C. C. & Lemon, S. M. 1996. Specific interaction of glyceraldehyde 3-phosphate dehyrogenase with the 5'-nontranslated RNA of hepatits A virus. J. Biol. Chem. **271**:14134-14142.

de Sena, J. & Mandel, B. 1977. Studies on the *in vitro* uncoating of poliovirus. Characteristics of the membrane-modified particle. Virology **78**:554-566.

Skern, T., Sommergruber, W., Blaas, D., Gruendler, P., Fraundorfer, F., Pieler, C., Fogy, I. & Kuechler, E. 1985. Human rhinovirus 2: complete nucleotide sequence and proteolytic signals in the capsid protein region. Nuc. Acids Res. **13**:2111-2126.

Skinner, M. A., Racaniello, V. R., Dunn, G., Cooper, J., Minor, P. D. & Almond, J. W. 1989. New model for the secondary structure of the 5' non-coding RNA of poliovirus is supported by biochemical and genetic data that also show that RNA secondary structure is important in neurovirulence. J. Mol. Biol. **207**:379-392.

Smith, T. J., Chase, E. S., Schmidt, T. J., Olson, N. H. & Baker, T. S. 1996. Neutralizing antibody to human rhinovirus 14 penetrates the receptor-binding canyon. Nature **383**:350-354.

Sommergruber, W., Zorn, M., Blaas, D., Fessl, F., Volkmann, P., Maurer-Fogy, I., Pallai, P., Merluzzi, V., Matteo, M., Skern, T. & Kuechler, E. 1989. Polypeptide 2A of human rhinovirus type 2: Identification as a protease and characterisation by mutational analysis. Virology **169**:68-77.

Sommergruber, W., Ahorn, H., Zöphel, A., Maurer-Fogy, I., Fessl, F., Schnorrenberg, G., Liebig, H.-D., Blaas, D., Kuechler, E. & Skern, T. 1992. Cleavage specificity on synthetic peptide substrates of human rhinovirus 2 protease 2A. J. Biol Chem. **267**:22639-22644.

Sommergruber, W., Carsri, G., Fessl, F., Seipelt, J. & Skern, T. 1994. The 2A protease of human rhinovirus is a zinc containing enzyme. Virology **204**:815-818.

Stanway, G., Can, A. J., Hauptmann, R., Hughes, P., Clarke, L. D., Mountford, R. C., Minor, P. D., Schild, G. C. & Almond, J. W. 1983. The nucleotide sequences of poliovirus type 3 Leon 12a1b: Comparison with poliovirus type 1. Nuc. Acids Res. **11**:5629-5643.

Stanway, G., Hughes, P. J., Mountford, R. C., Reeve, P., Minor, P. D., Schild, G. C. & Almond, J. W. 1984a. Comparison of the complete nucleotide sequences of the genomes of the neurovirulent poliovirus P3/Leon/37 and its attenuated Sabin vaccine derivative P3/Leon/12a1b. Proc. Natl. Acad. Sci. USA **81**:1539-1534.

Stanway, G., Hughes, P. J., Mountford, R. C., Minor, P. D & Almond, J. W. 1984b. The complete sequence of a common cold virus: human rhinovirus 14. Nuc. Acids Res. **12**:7859-7875.

Stanway, G. 1990. Structure, function and evolution of picornaviruses. J. Gen. Virol. **71**:2483-2501.

Staunton, D. E., Merluzzi, V. J., Rothlein, R., Barton, R., Marlin, S. D. & Springer, T. A. 1989. A cell adhesion molecule, ICAM-1, is the major surface receptor for rhinoviruses. Cell **56**:849-853.

Stone, D. M., Almond, J. W., Brangwyn, J. K. & Belsham, G. J. 1993. *Trans* complementaion of cap-independent translation directed by poliovirus 5' noncoding region deletion mutants: Evidence for RNA-RNA interactions. J. Virol. **67**:6215-6223.

Sun, X.-H. & Baltimore, D. 1989. Human immunodeficiency virus tat-activated expression of poliovirus protein 2A inhibits mRNA translation. Proc. Natl. Acad. Sci. USA **86**:2143-2146.

Supanaranond, K., Takeda, N. & Yamazaki, S. 1992. The complete nucleotide sequence of a variant of coxsackievirus A24, an agent causing acute hemorrhagic conjunctivitis. Virus Genes **6**:149-158.

Svitkin, Y. V., Ginevskaya, V. A., Ugarova, T. Y. & Agol, V. I. 1978. A cell free model of the encephalomyocarditis virus-induced inhibition of host cell protein synthesis. Virology **87**:199-203.

Svitkin, Y. V., Pestova, T. V., Maslova S. V. & Agol V. I. 1988. Point mutations modify the response of poliovirus RNA to a translation initiation factor: A comparison of neurovirulent and attenuated strains. Virology **166**:394-404.

Svitkin Y. V., Cammack, N., Minor, P. D. & Almond, J. W. 1990. Translation defficiency of the Sabin type 3 poliovirus genome: Association with an attenuating mutation $C_{472}\rightarrow U$. Virology **175:**103-109.

Svitkin, Y. V., Meerovitch, K., Lee, H. S., Dholakia, J. N., Kenan, D. J., Agol, V. I. & Sonenberg, N. 1994. Internal translation initiation on poliovirus RNA: Further characterisation of La function in poliovirus translation *in vitro*. J. Virol. **68**:1544-1550.

Svitkin, Y. V., Ovchinnikov, L. P., Dreyfuss, G. & Sonenberg, N. 1996. General RNA binding proteins render translation cap dependent. EMBO J. **15**:7147-7155.

Takegami, T., Kuhn, R. J., Anderson, C. W. & Wimmer, E. 1983. Membrane-dependent uridylydalation of the genome linked protein VPg of poliovirus. Proc. Natl. Acad. Sci. USA **80**:7447-7451.

Tan, E. M. 1989. Antinuclear antibodies: diagnostic markers for autoimmune diseases and probes for cell biology. Adv. Immunol. **44**:93-151.

Tardy-Panit, M., Blondel, B., Martin, A., Tekaia, F., Horaud, F. & Delpeyroux, F. 1993. A mutation in the RNA polymerase of poliovirus type 1 contributes to attenuation in mice. J. Virol. **67**:4630-4638.

Toolan, H. W. 1954. Transplantable human neoplasms maintained in cortisone-treated laboratory animals: HS#1; HEp#1; HEp#2; HEp#3 and HEmbRh#1. Cancer Res. **14**:660-673.

Toyoda, H., Kohara, M., Kataoka, Y., Sugamuma, T., Omata, T., Imura, N. & Nomoto, A. 1984. Complete nucleotide sequences of all three poliovirus serotype genomes. J. Mol. Biol. **174**:561-585.

Toyoda, H., Nicklin, M. J. H., Murray, M. G., Anderson, C. W., Dunn, J. J., Studier, F. W. & Wimmer, E. 1986. A second virus-encoded proteinase involved in proteolytic processing of poliovirus polyprotein. Cell **45**:761-770.

Trono, D., Andino, R. & Baltimore, D. 1988. An RNA sequence of hundreds of nucleotides at the 5' end of poliovirus RNA is involoved in allowing viral protein synthesis. J. Virol. **62**:2291-2299.

Tsukiyama-Kohara, A., Iizuka, N., Kohara, M. & Nomoto, A. 1992. Internal ribosome entry site within hepatitis C virus RNA. J. Virol. **66**:1476-1483.

Vagner, S., Gensac, M.-C., Maret, A., Bayard, F., Amalric, F., Prats, H. & Prats, A.-C. 1995. Alternative translation of human fibroblast growth factor 2 mRNA occurs by internal entry of ribosomes. Mol. Cell. Biol. **15**:35-44.

Van der Werf, S., Bradley, J., Wimmer, W., Studier, F. W. & Dunn, J. J. 1986. Synthesis of infectious poliovirus RNA by purified T7 RNA polymerase. Proc. Natl. Acad. Sci. USA **83**:2330-2334.

Ventoso, I. & Carrasco, L. 1995. A poliovirus $2A^{PRO}$ mutant unable to cleave 3CD shows inefficient viral protein synthesis and transactivation defects. J. Virol. **69**:6280-6288.

Wang, C., Sarnow, P. & Siddiqui, A. 1993. Translation of human hepatitis C virus RNA in cultured cells is mediated by an internal ribosome-binding mechanism. J. Virol. **67**:3338-3344.

Ward, T., Pipkin, P. A., Clarkson, N. A., Stone, D. M., Minor, P. D. & Almond, J. W. 1994. Decay-accelerating factor CD55 is identified as the receptor for echovirus 7 using CELICS, a rapid immuno-focal cloning method. EMBO J. **13**:5070-5074.

Weeks-Levy, C., Tatem, J. M., DiMichele, S. J., Waterfield, W., Georgiu, A. F. & Mento, S. J. 1991. Identification and characterisation of a new base substitution in the vaccine strain of Sabin 3 poliovirus. Virology **185**:934-937.

Westrop, G. D., Wareham, K. A., Evans, D. M. A., Dunn, G., Minor, P. D., Magrath, D. I., Taffs, F., Marsden, S., Skinner, M. A., Schild, G. S. & Almond, J. W. 1989. Genetic basis of attenuation of the Sabin type 3 oral poliovirus vaccine. J. Virol. **63**:1338-1344.

WHO 1983. Requirements for poliomyelitis vaccine (oral) (Revised 1982). Tech Reps. Ser. **687**:107-174, Geneva: World Health Organization.

WHO 1995. CVI Forum. Number 9.

Wyckoff, E. E., Hershey, J. W. B. & Ehrenfeld, E. 1990. Eukaryotic initiation factor 3 is required for poliovirus 2A protease-induced cleavage of the p220 component of eukaryotic initiation factor 4F. Proc. Natl. Acad. Sci. USA **87**:9529-9533.

Wyckoff, E. E., Lloyd, R. E. & Ehrenfeld, E. 1992. Relationship of eukaryotic initiation factor 3 to poliovirus induced p220 cleavage activity. J. Virol. **66**:2943-2951.

Yu, S. F. & Lloyd, R. E. 1991. Identification of essential amino acid residues in the functional activity of poliovirus 2A protease. Virology **182**:615-625.

Yu, S. F., Benton, P., Bovee, M., Sessions, J. & Lloyd, R. E. 1995. Defective RNA replication by poliovirus mutants deficient in 2A protease cleavage activity. J. Virol. **69**:247-252.

Zhang, G., Wilsden, G., Knowles, N. J. & McCauley, J. W. 1993. Complete nucleotide sequence of a coxsackie B5 virus and its relationship to swine vesicular disease virus. J. Gen. Virol. **74**:845-853.

Ziegler, E. Borman, A. M., Deliat, F. G., Liebig, H.-D., Jugovic, D., Kean, K. M., Skern. T. & Kuechler, E. 1995a. Picornavirus 2A proteinase-mediated stimulation of internal initiation of translation is dependent on enzymatic activity and the cleavage products of cellular proteins. Virology **213**:549-557.

Ziegler, E., Borman, A. M., Kirchweger, R.,Skern, T. & Kean, K. M. 1995b. Foot-and-mouth disease virus Lb proteinase can stimulate rhinovirus and enterovirus IRES-driven translation and cleave several proteins of cellular and viral origin. J. Virol. **69**:3465-3474.

All references are reproduced from original papers, retaining spelling and grammer.